

Db 462 GCGCTGTACGGCAGCAACCCATCGCGCGCGGTGAGCTACTTACCCCTCGACCCGCTG 521  
 Qy 505 GATTTGCTGTTGGACGACAGGCAATTCGCGGTGATGATGAAAAAGGTTACAGAGCCGC 564  
 Db 522 GACATCATCAAGACGCGGAGGACGTCGCGCGCCCGCTGAAGCCGCGCTACGAGTCGGCC 581  
 Qy 565 AACCGCGAATGACAAATACACTCGGTTTCGGTGTGAGCAACGACCGCGTGATGCCGCT 624  
 Db 582 AGCCACTCTGTTGACCTCGGCCACCGTCGCGCGCGCGCCGACACTTCGACGGGCTG 641  
 Qy 625 TTGCTGTATTCGACGTCGCGGTATGAGACCGAAGCGCGGCGAGGCTGCTATCCG 684  
 Db 642 CTGCAATTATGTCGCGCAGGCGCCACGAGACCGAATCCAACGGCGGCGCCACGGGGCAC 701  
 Qy 685 GTAGAGGGTGTGGCAGCGGA 705  
 Db 702 CGGGGCTCTTCGGGACGACGA 722

RESULT 2  
 BZ558933 1155 bp DNA linear GSS 17-DEC-2002  
 LOCUS pac82-164\_1131.s1 pac82-164 Pseudomonas aeruginosa genomic clone  
 DEFINITION pac82-164\_1131, genomic survey sequence.  
 ACCESSION BZ558933  
 VERSION BZ558933.1 GI:27174323  
 KEYWORDS GSS.  
 SOURCE Pseudomonas aeruginosa  
 ORGANISM Pseudomonas aeruginosa

REFERENCE Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 Pseudomonadaceae; Pseudomonas.  
 AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,  
 Burns, J.L., Kaul, R. and Olsen, M.V.

TITLE Whole-Genome-Sequence variation among multiple isolates of  
 Pseudomonas aeruginosa library  
 JOURNAL J. Bacteriol. (2002) In press  
 COMMENT Contact: Chris K. Raymond  
 Genome Center  
 University of Washington  
 Box 352145, Seattle, WA 98105-2145, USA  
 Tel: 2062216954  
 Fax: 2066857244  
 Email: craymond@u.washington.edu  
 Class: shotgun.

FEATURES source  
 Location/Qualifiers  
 1..1155  
 /organism="Pseudomonas aeruginosa"  
 /mol\_type="genomic DNA"  
 /strain="2-164"  
 /db\_xref="taxon:287"  
 /clone="pac82-164\_1131"  
 /clone\_lib="pac82-164"  
 /note="clinical isolate 2-164 Whole genomic shotgun library."

ORIGIN  
 Query Match 2.7%; Score 63.8; DB 8; Length 1155;  
 Best Local Similarity 49.7%; Pred. No. 7.2e-08;  
 Matches 189; Conservative 0; Mismatches 188; Indels 3; Gaps 1;  
 Qy 268 GCGTTTGTGTCGGCGCGTGAAGCAACCGTGTGCGGTGTCGAGCATTCACGCGCGTGAGC 327  
 Db 233 GCGTCCCAAGCGCGGACAGACACCAACATCTGGGGCTGGCGGACCGCGGAGAC 292  
 Qy 328 CTGCTGTATTCGAGAGAAACTCATGTATGACGTTATGGCAATTCGAACAGCTCGCC 387  
 Db 293 GTGCCCTATACCTGTGACGCGCGCAGAGAGCGGCTTCGAGCGCTTACCAGCAGGCGACC 352  
 Qy 388 CTGCTCTACGCCCACTCTGTCGCCACATCGCAATCGGAAGCGCGCTGACTCTTTC 447  
 Db 353 GTGTTTCATCAACCGGAAATGATCAAGCGCATCGAGGTGGAGAGGACCGCACTCGG 412

Qy 448 AATACCGGTAGCGCGCATTTGGTGGCGCGGTGAATTACAAACCGCTCGAAGGACATGAT 507  
 Db 413 TTCACCGCAATGGCGGCTTCGGCGGCACCGTGCACATGGAGACCAAGGACGCGCGGAC 472  
 Qy 508 TTGCTGTTGGACGACGAGCAATTCGCGGTGATGATGAAAAAGGTTACAGAGCGCGCAAC 567  
 Db 473 CTGCTGCGGAGGCGCGACGTCGCGCGCATGCTCAAAGTACGGCTATCACTCCCAACGAC 532  
 Qy 568 CGCAATGACAAATACACTCGGT--TTCGGTGTGAGCAACGACCGCGTGATGCCGCT 624  
 Db 533 CAGCAGAAGATCTACTTCGGGCGCTGTGTTGGCGCCACGANNAGACCGCGCGTGCATGCC 592  
 Qy 625 TTGCTGTATTGCAACGTCG 644  
 Db 593 CTGCTCTATCTCAACGGTCG 612

RESULT 3  
 BZ572005/c 916 bp DNA linear GSS 17-DEC-2002  
 LOCUS msh2\_225.y1 msh Pseudomonas aeruginosa genomic clone msh2\_225,  
 DEFINITION genomic survey sequence.  
 ACCESSION BZ572005  
 VERSION BZ572005.1 GI:27207066  
 KEYWORDS GSS.  
 SOURCE Pseudomonas aeruginosa  
 ORGANISM Pseudomonas aeruginosa

REFERENCE Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 Pseudomonadaceae; Pseudomonas.  
 AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,  
 Burns, J.L., Kaul, R. and Olsen, M.V.

TITLE Whole-Genome-Sequence variation among multiple isolates of  
 Pseudomonas aeruginosa library  
 JOURNAL J. Bacteriol. (2002) In press  
 COMMENT Contact: Chris K. Raymond  
 Genome Center  
 University of Washington  
 Box 352145, Seattle, WA 98105-2145, USA  
 Tel: 2062216954  
 Fax: 2066857244  
 Email: craymond@u.washington.edu  
 Class: shotgun.

FEATURES source  
 Location/Qualifiers  
 1..916  
 /organism="Pseudomonas aeruginosa"  
 /mol\_type="genomic DNA"  
 /strain="MSH"  
 /db\_xref="taxon:287"  
 /clone="msh2\_225"  
 /clone\_lib="msh"  
 /note="Environmental isolate. Whole genomic shotgun library."

ORIGIN  
 Query Match 2.7%; Score 63.6; DB 8; Length 916;  
 Best Local Similarity 50.3%; Pred. No. 7.7e-08;  
 Matches 184; Conservative 0; Mismatches 179; Indels 3; Gaps 1;  
 Qy 284 GCGTGAAGGCAACCGTGTGCGGTGTCAGCATTCAGCGGTGAGCTGCTGATTCGGAAG 343  
 Db 504 GCGGCAAGCCACCAACATCTGGGCGCTGGGCGACCGCGAGGACGTGCCCTATACCTGG 445  
 Qy 344 AAACTCACTGTATGACGTTATGCAACTTCAACAGCTCGCGCTGTCTATCGACCCCG 403  
 Db 444 ACGGCGCGCAAGAGATCGGCTTCGAGCGCTACACAGGGGACCGTGTTCATCGAACCG 385  
 Qy 404 AACTCGTGGCAACATCGAAATTCGGAAGGGCGCTGACTCTTTTCAATACCGGTAGCGCG 463  
 Db 384 AAATGATCAAGCGCATCGAGGTGGAGAGGGACCGCACTCGGTGTTCCACCGCAATCGCG 325  
 Qy 464 CATTGGTGGCGCGGTGAATTACAAACCGCTCAAGGACATGATTTGCTGTTGGACACA 523

Result No.	Score	Query Match	Length	DB	ID	Description
1	65.6	2.8	897	8	BZ572021	BZ572021 msh2_233.
2	63.8	2.7	1155	8	BZ558933	BZ558933 pac62_164
C 3	63.6	2.7	916	8	BZ572005	BZ572005 msh2_225.
C 4	40.2	1.7	493	9	CC624398	CC624398 OGUU574TV
C 5	40.2	1.7	864	9	CG323747	CG323747 CG2AS83TH
C 6	40.2	1.7	839	8	BZ553204	BZ553204 pac61-60
7	39	1.6	704	7	CO435339	CO435339 10a2-t3.W
C 8	38.8	1.6	1101	9	CNS00K2P	AL077149 drosophila
C 9	38.6	1.6	768	7	CF880971	CF880971 trico083xd
10	38.6	1.6	776	8	CB550719	CB550719 pac61-60
C 11	38.6	1.6	823	6	CB907731	CB907731 trico083xd
C 12	38.6	1.6	830	6	CA767580	CA767580 AF53-Rpf
C 13	38.6	1.6	910	5	BX928138	BX928138 BX928168
C 14	38.4	1.6	439	9	CL883336	CL883336 abf60e12.
C 15	38	1.6	486	2	BB851478	BB851478 BB851478
C 16	38	1.6	2181	3	AK077440	AK077440 Mus muscu
17	37.8	1.6	460	4	BI936731	BI936731 dc35c02.y
18	37.8	1.6	856	8	BZ564907	BZ564907 pac62-164
C 19	37.6	1.6	513	5	BQ825715	BQ825715 1030129C1
20	37.4	1.6	740	7	CF870398	CF870398 trico22xl
21	37.4	1.6	807	6	CB900579	CB900579 trico22xl
22	37.2	1.6	614	7	CF767164	CF767164 CES003263
C 23	37.2	1.6	718	7	CF881736	CF881736 trico085xd
C 24	37.2	1.6	843	6	CB908388	CB908388 trico085xd

QY

45 TTCAATACCGGTAGCGGCGCA

	Query Match	2.8%	Score 65.6	DB 8	Length 897
	Best Local Similarity	50.3%	Pred. No. 1.8e-08		
	Matches 161	Conservative	0	Mismatches 160	Indels 0
QY	385	CGCTGCTGTATCGACCCCGAACTGTGCGCAA	CATCGAAATCGCGAAGGGCGCTGACTCT	444	
Db	402	CGCAACTACGTCTGCTCGGCATCGTAAAGCGCGTGGAAATCTCTTCGCGGCCCGCGCTCG		461	
QY	445	TTCAATACCGGTAGCGGCGCATTTGGGTGGCGGCGTGAATTACCAAAACCTCTGCAAGGACAT		504	

1858 CTGAATGTGACAAAAGTAGCGTCTTTTGTCTGAGGGCTGGAATTTGTTGCGCTCGCTG 1917  
1861 CTGAATGTGACAAAAGTAGCGTCTTTTGTCTGAGGGTGGAACTGTTGCGTCTGCTG 1920  
1918 GGTATTGCGAAAGCAAACTGTGCGGCGCAACAGAGCTGCTGTCACAGCGCGCGAAA 1977  
1921 GGTATTGCGAAAGCAAACTGTGCGGCGCAACAGAGCTGCTGTCACAGCGCGCTGAAA 1980  
1978 GTGATTGCGGCTGCTGACATACGAAAGCGAGCGAAATGAGGCTGTTCTCCGCGCTG 2037  
1981 GTGATTGCGGCTGCTGACATACGAAAGCGAGCGAAATGAGGCTGTTCTCCGCGCTG 2040  
2038 ACTTATCTGGGTGCGAAAAAGCGCAAGAGCGCGCAATACACCGTTTATGAAACCAAGGCG 2097  
2041 ACTTATCTAGGCGGAAAAAGGTCAAAAGCGCGCAATACACCGTTTATGAAACCAAGGCG 2100  
2098 CGGGGTACGCTTTGAGAAAAAGGTAAAGATTACCGTGGCTGAAAGTGGCGTTAT 2157  
2101 TGGGGTACGCTTTGAGAAAAAGGTAAAGATTACCGTGGCTGAAAGTGGCGTTAT 2160  
2158 GTGTTTGATATGACGGCTTCTACAACTGGCTTAAACCTGACTTTGCGTGCAGCGTA 2217  
2161 GTGTTTGATATGACGGCTTCTACAAACCGGCTTAAACCTGACTTTGCGTGCAGCGTG 2220  
2218 TATAATGTGTTCAACCGCAAAATACACCACTTGGGATTCCTGCGCGGTTTGTATAGCTAC 2277  
2221 TACAACCTGTTCAACCGCAAAATACACCACTTGGGATTCCTGCGCGGTTTATAGCTAC 2280  
2278 AGCACCAACAGCGGTGACGAGATGCAAGAGGCTTACAGCGCTACCGCGCTCAGGC 2337  
2281 AGCACCAACAGCGGTGACGAGATGCAAGAGGCTTACAGCGCTTACCGCGCTCAGGC 2340  
2338 CGTAATTACGCGTATCGCTGGAATGGA 2373  
2341 CGCAATTACGCGTATCGCTGGAATGGA 2376

## RESULT 11

US-09-252-991A-14258/c  
; Sequence 14258, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 14258  
; LENGTH: 1401  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-14258

Query Match 2.9%; Score 68.4; DB 4; Length 1401;  
Best Local Similarity 50.5%; Pred. No. 5.4e-11;  
Matches 193; Conservative 0; Mismatches 186; Indels 3; Gaps 1;  
268 GCGTTTGTGTCGGCGGCTGGAGCAACCGTGTGCGTGTGAGCAATTCAGCGCTGAGC 327  
885 GCGTCCGCAAGCGCGCGGACAGACCAACATCTGGCGGCTTGGCGACGCCGAGGAC 826  
328 CTGCTCTGATTTCGGAAGAAAACCTCCTGTATGACGTTATGGCACTTCAACAGCTCGCGC 387  
825 GTGCGCTATACCTTGACGCGGCGGAGAGCGGCTTACGCGCTTACGAGGGGAC 766  
388 CTGCTCTATCGACCCCGAATCTGTCGCAACATCGAAATCGGAAGGGCGCTGACTCTTTC 447

765 GTGTTTCATCGAACCGGAAATGATCAAGCGCATCGAGGTGGAGAGCGGACCTCGCTG 706  
448 ATATACGGTAGCGCGCATTTGGTGGCGCGTGAATTACCAAAACCTCTGCAAGGACATGAT 507  
705 TTCACCGGCAATGGCGGCTTGGCGGCGACCGTGCAATGAGACCAAGGACCGCGCGGAC 646  
508 TTGCTGTTGGACGACAGGCAATTCGGCGTGTGATGATAAAGCGTTTACAGCAGCGCGCAAC 567  
645 CTGCTGCGGAGGCGCGGACGTCGCGGCCATGCTCAAGTACGGCTATCACTCCAAACGAC 586  
568 CCGAATGACAAATPAC---ACTCGGTTTGGTGTGAGCAACGCGCTGATGCGCT 624  
585 CAGCAGAAAGATCTACTCCGCGCGCTGTTGCGTGCAGCGAAGACCGCGCGCTCGATGCC 526  
625 TTGCTGTATTCCAAACGTCGCG 646  
525 CTGCTCTATCTCAACGGTCGCG 504

## RESULT 12

US-09-252-991A-14416  
; Sequence 14416, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 14416  
; LENGTH: 1767  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-14416

Query Match 2.9%; Score 68.4; DB 4; Length 1767;  
Best Local Similarity 50.5%; Pred. No. 6.4e-11;  
Matches 193; Conservative 0; Mismatches 186; Indels 3; Gaps 1;  
268 GCGTTTGTGTCGGCGGCTGGAGCAACCGTGTGCGTGTGAGCAATTCAGCGCTGAGC 327  
602 GCGTCCGCAAGCGCGGCGGACAGACCAACATCTGGCGGCTTGGCGACGCCGAGGAC 661  
328 CTGCTCTGATTTCGGAAGAAAACCTCCTGTATGCACTGTATGCACTTCAAGCTCGCGC 387  
662 GTGCGCTATACCTTGACGCGCGGCGGAGAGCGGCTTCGAGCGCTTACGAGGGGAC 721  
388 CTGCTCTATCGACCCCGAATCTGTCGCGCAACATCGAAATTCGGAAGGGCGCTGACTCTTTC 447  
722 GTGTTTCATCGAACCGGAAATGATCAAGCGCATCGAGGTGGAGAGGAGCGCACTCGGTG 781  
448 ATACCGGTAGCGGCGCATTTGGTGGCGGCTGATTTACCAACCTTCAGGACATGAT 507  
782 TTCACCGCAATGGCGGCTTGGCGGCGACCGTGCACTGGAGACCAAGGACCGCGCGGAC 841  
508 TTGCTGTTGGACGACAGCAATTTCCGCGTGTGATGATAAAGCGTTTACAGCAGCGCGCAAC 567  
842 CTGCTGCGGAGGCGCGGACGTCGCGGCCATGCTCAAGTACGGCTATCACTCCAAACGAC 901  
568 CCGAATGACAAATPAC---ACTCGGTTTGGTGTGAGCAACGCGCTGATGCGCT 624  
902 CAGCAGAAAGATCTACTCCGCGCGCTGTTGCGTGCAGCGAAGACCGCGCGCTCGATGCC 961  
625 TTGCTGTATTCCAAACGTCGCG 646  
962 CTGCTCTATCTCAACGGTCGCG 983



661 AGCGGGGCGAGCGTGGCTATCCGCTAGAGGGTCTGGCAGCGGAGCAATATTCGTTGGT 720  
Db  
661 AGCGGGGCAAGCGTGTATCCGCTAGAGGGTCTGGTAGCGGCGAATATCCGTTGGT 720  
Qy  
721 TCGTACGCGGGTATCCCTGATCCGCTCAAAACAACATACCAAACTTCTTGGGTGAAGTT 780  
Db  
721 TCTGGCGCGGGTATTCCTGATCCGCTCCCAACAACATACCAAACTTCTTGGGTGAAGTT 780  
Qy  
781 GCTTATCAATCAACGACAAGCACCGCATCGGCCATCGTTTAAAGCGGCAGCGGGCAT 840  
Db  
781 GCTTATCAATCAACGACAAGCACCGCATCGGCCATCGTTTAAAGCGGCAGCGGGCAT 840  
Qy  
841 AATTACAGATTTGAAGAGTCTTATACCTAGACCGCTTCTTCTGGCGGCAAGCCGATGAC 900  
Db  
841 AATTACAGATTTGAAGAGTCTTATACCTAGACCGCTTCTTCTTATTTGGCGTGAAGCTGACGAT 900  
Qy  
901 GTAACAGACGGGGCAATGCGCAACCTCTTTTACGAATGAGACCGCTGATTTCAAAATTTGGCTG 960  
Db  
901 GTCAACAGACGGGGTAAACCAACCTCTTTTACGAATGAGACCGCGGAATCCGACCGGTTG 960  
Qy  
961 TCGTCTTTCAAGCGCGACTTCGATTATCAGACCAACCAAGTGGCGGGTTAAACACAAA 1020  
Db  
961 TCTATGTTAAAGCGGATGCGATTATCAAAAACCAAAAGTATCTGCGGTCAACTACAAA 1020  
Qy  
1021 GGCTGCTTCCCGACCGGATTTCCACCTGGACCGCAACTATAATCAGAAGGATTTGGAG 1080  
Db  
1021 GGTTCGTTCCCGACCGAATTTACACCATGCGGAACCGAGTACCAATAAAAGGAAGTTGGC 1080  
Qy  
1081 AATATATCAACCGGACGATGAGACCGCGATTTCAAGCTTTTACCTTTGCGTATGACACAGC 1140  
Db  
1081 GAAATCTATAACCGGACGATGAGTAACTTTCAAAAGCTATTAACGCTGCGTATGACACAGC 1140  
Qy  
1141 CAACGCTTGCACACT--GGGCGGCAACCATCGCTTGTGCTTAAACCTTCCGCCAGTGG 1197  
Db  
1141 CATCGTTTGCACACTCGGGGGGGGGACACCGCTTGTGCTTAAACCTTTCGCCGGGAG 1200  
Qy  
1198 CGTGAGTTTGAACACTTAAACCGGACGATTTATTTCTCAGCGAAAGATATCCCGTACT 1257  
Db  
1201 CGTGATTTTGAACACTTAAACCGGACGATTTATTTCTCAGCGGCGGCTTGTTCGAACC 1260  
Qy  
1258 ACCAGCTGATTTCAACACCGCTGAAACCACTAATTTATGTTTCTCACTGCTGATCAA 1317  
Db  
1261 ACCAACAGTATCCAGCATCCGCTGAAACCACTACGCTTCTCGCTGCTCGGACCAA 1320  
Qy  
1318 ATCCAAATGGAACGAGCTGTTACGAGCGCTGAGATATCCGTTACGATCATACCAAAATG 1377  
Db  
1321 ATCCAAATGGAACGAGCTGTTACGAGCGCTGAGATATCCGTTACGATCATACCAAAATG 1380  
Qy  
1378 AGCCTTCAGGAATGAATGCGAGTGTCTATGCTTGTGACAAACACCGCTTCGACCCAA 1437  
Db  
1381 AGCCTTCAGGAATGAATGCGAGTGTCTATGCTTGTGACAAACACCGCTTCGACCCAA 1440  
Qy  
1438 ACTTATAAGGCTGAGCGGATTTGCGTTTGGCGGCGCAACTGAATCAGGCTTTGGCAT 1497  
Db  
1441 ACTTATAAGGCTGAGCGGATTTGCGTTTGGCGGCGCACTGAGCGCAACATGCGGT 1500  
Qy  
1498 GTCGGTTACGACATTTCTCGGCTACGCTGCCCAATGCTGCCGAAGTGTATTTCACT 1557  
Db  
1501 GTGGGTTACGATGTACCTCAGTTTCCCGTGGCGAATGCGTGTGAAGTGTATTTCACT 1560  
Qy  
1558 TACAACACCGGTTCCGGTAAATTTGGCTGCCCAATCCCAACCTGAAAGCGGAGCGACGACC 1617  
Db  
1561 TACAACACCGGTTCCGGCACTTTGGAAGCTTAATCTTAATTTGAAGCGAAGCGACGACC 1620  
Qy  
1618 ACCCAACCTGCTCTCAAGCGCGGCGGCAAAAGGATCTTTGGATGCCAACTGTAT 1677  
Db  
1621 ACCCAACCTGCTCTCAAGCGGCGGCGGCAAAAGGATCTTTGGATGCCAACTGTAT 1680  
Qy  
1678 CAABAACATTCGCAACTTCTTGTCTGAAGAGCGAGCTCAACACGAGCGGCGATGTC 1737  
Db  
1681 CAAGAATATTCGCAACTTCTTGTCTGGAAGAGCGAGATCTGACTGTGAGCGGCAACCC 1740  
Qy  
1738 GGCTGTACTCAGATGAATTTACTACTACGCTATGTGTAGCAATCTTATTTCCGAAAAACCG 1797

1741 GGCTGTACTGAGGAGTCTTACTACTATAGATCAGCGACCCCTACAAAGAAAACCTG 1800  
Qy  
1798 GAATGCGAGATGCAAAATATCGATAAGGCCCAATCCGTTGCTTGTAGCTGACAGGCCGT 1857  
Db  
1801 GATTGCGAGATGAAAAATATCGACAAGGCCAAGATCCGCGTATCGAGTTGACAGGCCGT 1860  
Qy  
1858 CTGAATGTGACAAAAGTAGCGTCTTTTCTGCTGAGGCTGGAATTTGCTCGCTCGCTG 1917  
Db  
1861 CTGAATGTGACAAAAGTAGCGTCTTTTCTGCTGAGGTTGGAATTTGCTCGCTCGCTG 1920  
Qy  
1918 GGTATGCAAAAGCAAACTGTGCGGCGCAACACAGCTGCTGTCCACACAGCGCGCAAA 1977  
Db  
1921 GGTATGCAAAAGCAAACTGTGCGGCGCAACACAGCTGCTGTCCACACAGCGCGTGA 1980  
Qy  
1978 GTGATTCGCGTGTGCTACGACTACGAAGCCGCGCAAAAGTGGGTTGTTCTCCCGCTG 2037  
Db  
1981 GTGATTCGCGTGTGCTACGACTATGAAGTCCGAGCGAAAAATGGGCGTATTTCTCCCGCTG 2040  
Qy  
2038 ACTTATCTGGGTGCGAAAAAGCGCAAGACGCGCAATACACCGTTTATGAAAAACAAGGCG 2097  
Db  
2041 ACCTATCTAGGCGCAAAAGGTCAAGACGCGCAATACACCGTTTATGAAAAACAAGGCG 2100  
Qy  
2098 CGGGTAGCGCTTTGCGAAAAAGGTAAAGATTAACCGTGGCTGAAACAAGTCCGCTTAT 2157  
Db  
2101 TGGGTACGCTTTGCGAAAAAGGTAAAGATTAACCGTGGCTGAAACAAGTCCGCTTAT 2160  
Qy  
2158 GTGTTGATGATGACGGCTTCTCAAACTGGCTTAAACCTGACTTTGGTGCAGCGCTA 2217  
Db  
2161 GTGTTGATGATGACGGCTTCTCAAACTGGCTTAAACCTGACTTTGGTGCAGCGCTG 2220  
Qy  
2218 TATAATGTGTTCAACCGCAAAATACACCACTTGGGATTCCTCGCGGTTTGTATAGCTAC 2277  
Db  
2221 TACAACCTGTTCAACCGCAAAATACACCACTTGGGATTCCTCGCGGTTTGTATAGCTAC 2280  
Qy  
2278 AGCACCAACGCGGTGCGACCGAGATGCGAAAGCTTTAGACCGCTACCGCGCTCAGGC 2337  
Db  
2281 AGCACCAACGCGGTGCGACCGAGATGCGAAAGCTTTAGACCGCTACCGCGCCCGCAGC 2340  
Qy  
2338 CGTAATGACGCGTATCGCTGATGGAAGTTTGA 2373  
Db  
2341 CGCAATGACGCGTATCGCTGATGGAAGTTTGA 2376

RESULT 10  
US-08-537-361E-3  
; Sequence 3, Application US/08537361E  
; Patent No. 6121037  
; GENERAL INFORMATION:  
; APPLICANT: Stojiljkovic, Igor  
; APPLICANT: So, Magdalene  
; APPLICANT: Hwa, Vivian  
; APPLICANT: Heffron, Fred  
; APPLICANT: Nassif, Xavier  
; TITLE OF INVENTION: No. 6121037el Bacterial Hemoglobin Receptor  
; TITLE OF INVENTION: Genes and Uses  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive, 32nd Floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/537,361E  
; FILING DATE: 02-OCT-1995  
; CLASSIFICATION: 536











QY	1495	CATGTCGGTTACGACATTA	CTTCGGGCTACCGTGTCCC	CAATCGCTCCGAAGTGAT	TTTC	1554	
DB	1971	CGTGTCCGTTACGACATTA	CTTCGGGCTACCGTGTCCC	CAATCGCTCCGAAGTGAT	TTTC	2030	
QY	1555	ACTTACAACACACGGTTCGG	GTAAATGGCTGCCCAATCC	GAACTCGAAGACCGAGCGC	GAC	1614	
DB	2031	ACTTACAACACACGGTTCGG	GTAAATGGCTGCCCAATCC	GAACTCGAAGACCGAGCGC	GAC	2090	
QY	1615	ACCACCCACACCCCTGTCT	CTGCAAGCGCGCAGCGAAA	AGAGTACTTTGGATGCCAAC	CTG	1674	
DB	2091	ACCACCCACACCCCTGTCT	CTGCAAGCGCGCAGCGAAA	AGAGTACTTTGGATGCCAAC	CTG	2150	
QY	1675	TATCAAAA	CAATTA	CCGCAACTTCCTGTAAGAG	CAGAAGCTGACCA	CCAGCGCGCAT	1734
DB	2151	TATCAAA	GCAATTA	CCGCAACTTCCTGTAAGAG	CAGAAGCTGACCA	CCAGCGCGCAT	2210
QY	1735	GTCCGCTGTACTCAGATGA	ATTA	CTACTCTACGTTATGTTAG	CAATCCTTAATCCGAAAA	1794	
DB	2211	GTCACTGTACTCAGATGA	ATTA	CTACTCTACGTTATGTTAG	CAATCCTTAATCCGAAAA	2270	
QY	1795	CCGGAATGGCAGATGCAAAA	TATCGATAAGGCCCGGAAT	CCTGTGGTCTTTGAGCTG	CACAGGC	1854	
DB	2271	CTGGAATGGCAGATGCAAAA	TATCGATAAGGCCCGGAAT	CCTGTGGTCTTTGAGCTG	CACAGGC	2330	
QY	1855	CGTCTGAATGTGACAAAGTA	GCGTCTTTTGTCTTGAGGG	CTGGAAATTTGTTCCGGCTCG	1914		
DB	2331	CGTCTGAATGTGACAAAGTA	GCGTCTTTTGTCTTGAGGG	CTGGAAATTTGTTCCGGCTCG	2390		
QY	1915	CTGGGTTATCGGAAAAAGCA	AACTGTGCGGCGACAAACAG	CTGCTGCCACACAGCCGCGC	1974		
DB	2391	CTGGGTTATCGGAAAAAGCA	AACTGTGCGGCGACAAACAG	CTGCTGCCACACAGCCGCGT	2450		
QY	1975	AAAAGTGATTCGCGTGTGCA	CTACGAAAGCCCGACGAAAA	ATGGGGTGTGTTCTCCCGC	2034		
DB	2451	AAAAGTGATTCGCGTGTGCA	CTACGAAAGCCCGACGAAAA	ATGGGGTGTGTTCTCCCGC	2510		
QY	2035	CTGACTTATCTGGGTGCGAAA	AGGCCAAAGACGCGCAATAC	ACCGTTTATGAAAAACAAG	2094		
DB	2511	CTGACCTATCTGGGCGCGAAA	AGGGTCAAGACGCGCAATAC	ACCGTTTATGAAAAACAAG	2570		
QY	2095	GGCCGGGTACGCTTTGCA	GAAAAAGTAAAGATTAC	CCGTGGCTGAAACAAGTCGGCT	2154		
DB	2571	GGCTGGGTACGCTTTGCA	GAAAAAGTAAAGATTAC	CCGTGGCTGAAACAAGTCGGCT	2630		
QY	2155	TATGTGTTTGATATGTACG	GCTTCTACAAACTGGCTAAAA	ACCTGACTTTTCGTCGACAGC	2214		
DB	2631	TATGTGTTTGATATGTACG	GCTTCTACAAACTGGCTAAAA	ACCTGACTTTTCGTCGACAGC	2690		
QY	2215	GTATATAATGTGTTCAACCG	CAATAACACACTTTGGGAT	TTCCTGCGCGGTTTGTATAGC	2274		
DB	2691	GTATATAATGTGTTCAACCG	CAATAACACACTTTGGGAT	TTCCTGCGCGGCTTGTATAGC	2750		
QY	2275	TACAGCACCAACCGCGT	TCAGCCGAGATGGCAAAAG	CGTTAGACCGCTTACCGCGCCTCA	2334		
DB	2751	TACAGCACCAACCGCGT	TCAGCCGAGATGGCAAAAG	CGTTAGACCGCTTACCGCGCCTCA	2810		
QY	2335	GGCCGTAAATACGCCGTAT	CGCTGGATGGAAAGTTTTGA	2373			
DB	2811	AGCCGTAAATACGCCGTAT	CGCTGGATGGAAAGTTTTAA	2849			

## RESULT 7

US-08-537-361E-5

Sequence 5, Application US/08537361E

; Patent No. 6121037

GENERAL INFORMATION:

APPLICANT: Stojiljkovic, Igor

APPLICANT: So, Magdalene

APPLICANT: Hwa, Vivian

APPLICANT: Hefron, Fred

APPLICANT: Nassif, Xavier

**TITLE OF INVENTION:** No. 6121037el Bacterial Hemoglobin Receptor

TITLE OF INVENTION: Genes and Uses  
 NUMBER OF SEQUENCES: 14  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: McDonnell Boeheim Hulbert & Berghoff  
 STREET: 300 South Wacker Drive, 32nd Floor  
 CITY: Chicago  
 STATE: Illinois  
 COUNTRY: USA  
 ZIP: 60606  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/537,361E  
 FILING DATE: 02-OCT-1995  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: No. 6121037nan, Kevin E  
 REGISTRATION NUMBER: 35,303  
 REFERENCE/DOCKET NUMBER: 94,784-A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312-913-0001  
 TELEFAX: 312-913-0002  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2379 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..2376  
 US-08-537-361E-5

Query Match 83.4%; Score 1983.8; DB 3; Length 2379;

Query Match 90.28; Pred. No. 0;  
Best Local Similarity 90.28; Pred. No. 0;

Matches 2146; Conservative 0; Mismatches 227; Indels 6; Gaps 2;

1 ATGAAACCATTAACATGCTTCCCTATTGCCGCGCTGGTCGGCAGTATTTTCGGCAATCCG 60

0  
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99

61 CTTCTCCGACCGGATGTAAGCTGCAACCGAACCACACCGTTAAGCAGAGATAAAGAA 12

61 GTC TGG CAG CGG ATG AAG CTG CAA CCG AAT CCA CAC CCG TAT AAG CAG AGA TAT AAG AA 12

0  
1  
2  
3  
4  
5  
6  
7  
8  
9  
A  
B  
C  
D  
E  
F  
G  
H  
I  
J  
K  
L  
M  
N  
O  
P  
Q  
R  
S  
T  
U  
V  
W  
X  
Y  
Z  
[  
\  
]  
^  
\_  
`  
a  
b  
c  
d  
e  
f  
g  
h  
i  
j  
k  
l  
m  
n  
o  
p  
q  
r  
s  
t  
u  
v  
w  
x  
y  
z  
{  
|  
}  
~  
\_

Qy 121 GTGCGCGTTAAAGACCAACAGCTTAATGCGCCCTGCAACCGTGGAAACGTGTCAACCTCGGCCGC 18

QY 181 ATTCAACAGGAAATGATACGCGACAAACAAGACTTGGTGGCTTACTCCACCGACGTCGGC 24

.....

QY 241 TTGAGCGATAGCGGCCGCATCAAAAGGCTTTGCTGTGCGCGCGTGGAGGCAACCGT 30

DD 241 TIGAGCGACAGCGCGCCGCCATCAATTAAGGCATTAGCCGTTCCGCGCGTGGATGGCGATACCCCGT

301 GTCCGGTGTCA GCATTGACGGCGTGAGCCTGCCTGATTCGGAAGAAAACTCAGCTGTATGCA 36

Db 301 GTCGCGGTGAGCATAGACGGCGTAAACCTGCC<sup>1</sup>TGATTCCGAAGAAAC<sup>1</sup>TGCT<sup>1</sup>GTACGCC 36

2. 3. 5. 7. 9. 11. 13. 15. 17. 19. 21. 23. 25. 27. 29. 31. 33. 35. 37. 39. 41. 43. 45. 47. 49. 51. 53. 55. 57. 59. 61. 63. 65. 67. 69. 71. 73. 75. 77. 79. 81. 83. 85. 87. 89. 91. 93. 95. 97. 99. 101. 103. 105. 107. 109. 111. 113. 115. 117. 119. 121. 123. 125. 127. 129. 131. 133. 135. 137. 139. 141. 143. 145. 147. 149. 151. 153. 155. 157. 159. 161. 163. 165. 167. 169. 171. 173. 175. 177. 179. 181. 183. 185. 187. 189. 191. 193. 195. 197. 199. 201. 203. 205. 207. 209. 211. 213. 215. 217. 219. 221. 223. 225. 227. 229. 231. 233. 235. 237. 239. 241. 243. 245. 247. 249. 251. 253. 255. 257. 259. 261. 263. 265. 267. 269. 271. 273. 275. 277. 279. 281. 283. 285. 287. 289. 291. 293. 295. 297. 299. 301. 303. 305. 307. 309. 311. 313. 315. 317. 319. 321. 323. 325. 327. 329. 331. 333. 335. 337. 339. 341. 343. 345. 347. 349. 351. 353. 355. 357. 359. 361. 363. 365. 367. 369. 371. 373. 375. 377. 379. 381. 383. 385. 387. 389. 391. 393. 395. 397. 399. 401. 403. 405. 407. 409. 411. 413. 415. 417. 419. 421. 423. 425. 427. 429. 431. 433. 435. 437. 439. 441. 443. 445. 447. 449. 451. 453. 455. 457. 459. 461. 463. 465. 467. 469. 471. 473. 475. 477. 479. 481. 483. 485. 487. 489. 491. 493. 495. 497. 499. 501. 503. 505. 507. 509. 511. 513. 515. 517. 519. 521. 523. 525. 527. 529. 531. 533. 535. 537. 539. 541. 543. 545. 547. 549. 551. 553. 555. 557. 559. 561. 563. 565. 567. 569. 571. 573. 575. 577. 579. 581. 583. 585. 587. 589. 591. 593. 595. 597. 599. 601. 603. 605. 607. 609. 611. 613. 615. 617. 619. 621. 623. 625. 627. 629. 631. 633. 635. 637. 639. 641. 643. 645. 647. 649. 651. 653. 655. 657. 659. 661. 663. 665. 667. 669. 671. 673. 675. 677. 679. 681. 683. 685. 687. 689. 691. 693. 695. 697. 699. 701. 703. 705. 707. 709. 711. 713. 715. 717. 719. 721. 723. 725. 727. 729. 731. 733. 735. 737. 739. 741. 743. 745. 747. 749. 751. 753. 755. 757. 759. 761. 763. 765. 767. 769. 771. 773. 775. 777. 779. 781. 783. 785. 787. 789. 791. 793. 795. 797. 799. 801. 803. 805. 807. 809. 811. 813. 815. 817. 819. 821. 823. 825. 827. 829. 831. 833. 835. 837. 839. 841. 843. 845. 847. 849. 851. 853. 855. 857. 859. 861. 863. 865. 867. 869. 871. 873. 875. 877. 879. 881. 883. 885. 887. 889. 891. 893. 895. 897. 899. 901. 903. 905. 907. 909. 911. 913. 915. 917. 919. 921. 923. 925. 927. 929. 931. 933. 935. 937. 939. 941. 943. 945. 947. 949. 951. 953. 955. 957. 959. 961. 963. 965. 967. 969. 971. 973. 975. 977. 979. 981. 983. 985. 987. 989. 991. 993. 995. 997. 999. 1001. 1003. 1005. 1007. 1009. 1011. 1013. 1015. 1017. 1019. 1021. 1023. 1025. 1027. 1029. 1031. 1033. 1035. 1037. 1039. 1041. 1043. 1045. 1047. 1049. 1051. 1053. 1055. 1057. 1059. 1061. 1063. 1065. 1067. 1069. 1071. 1073. 1075. 1077. 1079. 1081. 1083. 1085. 1087. 1089. 1091. 1093. 1095. 1097. 1099. 1101. 1103. 1105. 1107. 1109. 1111. 1113. 1115. 1117. 1119. 1121. 1123. 1125. 1127. 1129. 1131. 1133. 1135. 1137. 1139. 1141. 1143. 1145. 1147. 1149. 1151. 1153. 1155. 1157. 1159. 1161. 1163. 1165. 1167. 1169. 1171. 1173. 1175. 1177. 1179. 1181. 1183. 1185. 1187. 1189. 1191. 1193. 1195. 1197. 1199. 1201. 1203. 1205. 1207. 1209. 1211. 1213. 1215. 1217. 1219. 1221. 1223. 1225. 1227. 1229. 1231. 1233. 1235. 1237. 1239. 1241. 1243. 1245. 1247. 1249. 1251. 1253. 1255. 1257. 1259. 1261. 1263. 1265. 1267. 1269. 1271. 1273. 1275. 1277. 1279. 1281. 1283. 1285. 1287. 1289. 1291. 1293. 1295. 1297. 1299. 1301. 1303. 1305. 1307. 1309. 1311. 1313. 1315. 1317. 1319. 1321. 1323. 1325. 1327. 1329. 1331. 1333. 1335. 1337. 1339. 1341. 1343. 1345. 1347. 1349. 1351. 1353. 1355. 1357. 1359. 1361. 1363. 1365. 1367. 1369. 1371. 1373. 1375. 1377. 1379. 1381. 1383. 1385. 1387. 1389. 1391. 1393. 1395. 1397. 1399. 1401. 1403. 1405. 1407. 1409. 1411. 1413. 1415. 1417. 1419. 1421. 1423. 1425. 1427. 1429. 1431. 1433. 1435. 1437. 1439. 1441. 1443. 1445. 1447. 1449. 1451. 1453. 1455. 1457. 1459. 1461. 1463. 1465. 1467. 1469. 1471. 1473. 1475. 1477. 1479. 1481. 1483. 1485. 1487. 1489. 1491. 1493. 1495. 1497. 1499. 1501. 1503. 1505. 1507. 1509. 1511. 1513. 1515. 1517. 1519. 1521. 1523. 1525. 1527. 1529. 1531. 1533. 1535. 1537. 1539. 1541. 1543. 1545. 1547. 1549.

Db 361 CGTTATGGCAACTTCAACAGCTCGCGTCTGTCTATCGACCCCGAACTCGTGGCGCAACATC 42

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

APPLICANT: Stojiljkovic, Igor  
 APPLICANT: So, Magdalena  
 APPLICANT: Hwa, Vivian  
 APPLICANT: Heffron, Fred  
 APPLICANT: Nassif, Xavier

TITLE OF INVENTION: No. 6121037el Bacterial Hemoglobin Receptor  
 TITLE OF INVENTION: Genes and Uses  
 NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff

STREET: 300 South Wacker Drive, 32nd Floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/537,361E

FILING DATE: 02-OCT-1995

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: No. 6121037nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 94,784-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-913-0001

TELEFAX: 312-913-0002

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3319 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 471..2848

US-08-537-361E-1

Query Match 83.5%; Score 1985.4; DB 3; Length 3319;  
 Best Local Similarity 90.2%; Pred. No. 0;  
 Matches 2147; Conservative 0; Mismatches 226; Indels 6; Gaps 2;

QY 1 ATGAAACCATTTACATGCTTCTTATTTGCGCGCTGCTGGCGAGTATTTTCGGCAATCCG 60  
 DB 471 ATGAAACCATTTACAAATGCTCCCTATCGCGCGCTGCTGGCGAGTATTTTCGGCAATCCG 530  
 QY 61 GTCTTGGCAGGATGAGCTGCACCAACCAACACACCCGTTAAAGCAGAGATAAGAA 120  
 DB 531 GTCTTGGCAGGATGAGCTGCACCAACCAACCAACACCCGTTAAAGCAGAGATAAGAA 590  
 QY 121 GTGCGCGTTAAAGACCAACAGCTTAATCGCGCTGCAACCGTGGAACTGTCAACCTCGGCGCG 180  
 DB 591 GTGCGCGTTAAAGACCAACAGCTTAATCGCGCTGCAACCGTGGAACTGTCAACCTTAACCGT 650  
 QY 181 ATTCAACAGGAATGATACCGGCAACAAAGACTGTGGTGGTGTACTCCACCGAGCTGGC 240  
 DB 651 ATCAAAACAAGAAATGATACCGGCAACAAAGACTGTGGTGGTGTACTCCACCGAGTGTGGC 710  
 QY 241 TTGAGCGATACCGCGCCCATCAAAAGGCTTTGCTGTGCGCGCGTGGAGGCAACCGT 300  
 DB 711 TTGAGCGATACCGCGCCCATCAAAAGGCTTTGCTGTGCGCGCGTGGAGGCAACCGT 770  
 QY 301 GTGCGGTGTGAGCTTGAACCGGCTGAGCTGCTGATTCGGAAGAAAACCTCACTGTATGCA 360  
 DB 771 GTGCGGTGTGAGCTTGAACCGGCTGAGCTGCTGATTCGGAAGAAAACCTCACTGTATGCA 830  
 QY 361 CGTTATGGCAACTTCAACAGCTCGCGCTGTCTATCGACCCCGAACTCGTGGCGCAACATC 420

DB 831 CGTTATGGCAACTTCAACAGCTCGCGTCTGTCTATCGACCCCGAACTCTGTGCGCAACATC 890  
 QY 421 GAAATGCGGAAGCGCGCTGACTCTTTCAATACCGGTAGCGCGCGCATTTGGGTGGCGGCTG 480  
 DB 891 GACATGTAAGAGGCGGCGACTCTTTCAATACCGCGAGCGCGCTTGGCGGCGGCTG 950  
 QY 481 AATTACCACACCTTGAAGGACATGATTTGCTGTGTGACGACAGGCAATTTCCGCGTGATG 540  
 DB 951 AATTACCACACCTTGAAGGACGCTGACTTACTGTGCTGAACGGCAGTTTCCGCGTGATG 1010  
 QY 541 ATGAAAAACGTTACAGACCGCGCAACCGGCAATGAGCAATATACACTCGGTTTCGGTGTG 600  
 DB 1011 ATGAAAAACGTTACAGACCGCGTAACCGTGAATGGAACAATACCTTCGCTTTCGGCGTG 1070  
 QY 601 AGCAACGACCGCTGGATGCGCTTTGCTGCTATTTCGCAACGCTCGCGTCAATGAGACGAA 660  
 DB 1071 AGCAACGACCGCTGGATGCGCTTTGCTGCTATTTCGCAACGCGCGCGCAATGAACTGAA 1130  
 QY 661 AGCGCGGCGAGCGTGCTATCCGCTAGAGGGTCTGGCAGCGGAGCAATTTATCCGTGGT 720  
 DB 1131 AGCGCGGCGAAGCGTGTATTCCGCTAGAGGGTCTGGTAGCGAGCAATATCCGTGGT 1190  
 QY 721 TCGTACGCGGTATCCCTGATCCGTCGAACACAAATACACAACTTCTTTGGGTAAAGATT 780  
 DB 1191 TCTGCGCGCGGTATTCCTGATCCGTCCTCCCAACACAAATACCAACAGCTTCTTTGGGTAA 1250  
 QY 781 GCTTATCAATCAACGACGACCGCATCGCGCCATCGTTTAAACGCGCAGCAGGGGCAT 840  
 DB 1251 GCTTATCAATCAACGACGACCGCATCGCGCCATCGCTCAACCGTTCAGCAGGGGCAT 1310  
 QY 841 AATTACACGATTGAAGAGTCTTATAACCTGACCGCTTCTTCTGCGCGCAAGCGATGAC 900  
 DB 1311 AATTACACGATTGAAGAGTCTTATAACCTGACCGCTTCTTCTTATTTGGGTGAAGTGA 1370  
 QY 901 GTAAACAGACGCGCAATGCCAACCCTTTTACGAATGAGCGCTGATTCAAATGGCTG 960  
 DB 1371 GTCAACAGACGCGGTAAACCAACCTCTTTTACGAATGAGCGCGGAATCCGACCGGTTG 1430  
 QY 961 TCGTCTTTGAAAGCGGACTTCGATTATCAGACAAACAAAGTGGCGGCTTAAACACAAA 1020  
 DB 1431 TCTATGTTAAAGCGGATGCTGATTATCAAAAACCAAGTATCTGCGGTCAACTACAA 1490  
 QY 1021 GGCTCGTTCGCGA --- CGGATTATTCACCTGGACGCGCAACTATATATCAGAGAGATTG 1077  
 DB 1491 GGTTCGTTCCGATAGAGGATCTTCCACCTTGACACGTAACATAATCAAAAGGACTTG 1550  
 QY 1078 GAGATATATACAAACGCGACGATGGACACCGGATTCAAACGTTTACTTTGCGTATGGAC 1137  
 DB 1551 GATGAAATCTACAAACGCGAGTATGGATACCGCTTCAACCGCATTAACCTGCGTTGGAC 1610  
 QY 1138 AGCCAAACCGTTGCAACT --- GGGCGCGCAACATCGCTTGTGCTTAAACCTTTTCGCCAGT 1194  
 DB 1611 AGCCAAACCGTTGCAACT --- GGGCGCGCGGCGGCGACACCGCTGCTTTTAAACCTTTCCGCGAGC 1670  
 QY 1195 CGCGGTGAGTTGAAAACCTTAAACCGCGACGATTATTTACTTTCAGCGAAAGAGTATCCCGT 1254  
 DB 1671 CGCGGTGATTTTGAACCTTAAACCGCGACGATTATTTACTTTCAGCGCGCGTGTGTTGTCGA 1730  
 QY 1255 ACTACAGCTGATTCAACACCCCGTGAAAACCACTAAATTTATGTTTCTCACTGCTGAT 1314  
 DB 1731 ACCACGACGAGTATTCAGCATCCGTTGAAAACCAACCACTACCGTTTCTCACTGCTGAC 1790  
 QY 1315 CAAATCAATGGAACGACGCTTTCAGCAGCGCTGAGATATCCGTTACGATCATACCAAA 1374  
 DB 1791 CAAATTCATGGAACGACGCTTTCAGTACCGCGCAGGATATCCGTACGATCATACCAAA 1850  
 QY 1375 ATGACGCTCAGGAATTGAATCCGAGTGTGTCATGTTGTGCAAAAAACCGCTTCAGGCC 1434  
 DB 1851 ATGACGCTCAGGAATTGAATCCGAGTGTGTCATGTTGTGCAAAAAACCGCTTCAGGCC 1910  
 QY 1435 AATACCTTATAAGGCTGGAGCGGATTTGCGGTTTGGCGCGCAACTGAATCAGGCTTGG 1494  
 DB 1911 AACACTTATAAGGCTGGAGCGGTTTGTGCGCTTGGCGCGCAACTGAATCAGGCTTGG 1970

301 GTCTGTTGAGCAGTGAACGGCGTGAAGCTGCTGATTCGGAAGAAATCACTGATGCA 360  
Db  
770 GTCTGCGGTGAGATAGACGGCGTAAACCTGCTGATTCGGAAGAAATCACTGATGCGC 829  
QY  
361 CGTTATGGCAACTCAACAGCTGGCGCTGTCTATCGACCCCGAACTCTGTGGCGCAACATC 420  
Db  
830 CGTTATGGCAACTCAACAGCTGGCGCTGTCTATCGACCCCGAACTCTGTGGCGCAACATC 889  
QY  
421 GAAATCGCGAAGCGGCTGACTCTTTCAATACCGGTAGCGGCGCAATGGTGGCGCGCGTG 480  
Db  
890 GACATCGTAAAGGGCGGACTCTTTCAATACCGGAGCGGCGCTTGGGCGCGCGGTG 949  
QY  
481 AATTACCAAACTGCAAGGACATGATTTGCTTTGGACGACGAGCAATTCGCGGTGATG 540  
Db  
950 AATTACCAAACTGCAAGGACATGATTTGCTTTGGCTGAAACGCGAGTTGCGGCGGTG 1009  
QY  
541 ATGAAACCGGTACAGACGCGCAACCGGAAATGGAACAAATACACTCGTTTCGGGTG 600  
Db  
1010 ATGAAACCGGTACAGACGCGCAACCGGAAATGGAACAAATACACTCGTTTCGGGTG 1069  
QY  
601 AGCAACGCGGCTGATGCGCTTTGCTGATTCGCAACGTCGCGGTCACTGAGACCGAA 660  
Db  
1070 AGCAACGCGGCTGATGCGCTTTGCTGATTCGCAACGTCGCGGTCACTGAGACCGAA 1129  
QY  
661 AGCGCGGCGAGCTGCTATCCGCTAGAGGTGCTGGCAGCGGAGCAATTAATCCGCTGT 720  
Db  
1130 AGCGCGGCGAGCTGCTATCCGCTAGAGGTGCTGGCAGCGGAGCAATTAATCCGCTGT 1189  
QY  
721 TCGTACGCGGTATCCGCTAGAGGTGCTGGCAGCGGAGCAATTAATCCGCTGT 780  
Db  
1190 TCGTACGCGGTATCCGCTAGAGGTGCTGGCAGCGGAGCAATTAATCCGCTGT 1249  
QY  
781 GCTTATCAATCAACGACGCGCATCGGCGCATCGTTTAAACGCGGAGCGGCGAT 840  
Db  
1250 GCTTATCAATCAACGACGCGCATCGGCGCATCGTTTAAACGCGGAGCGGCGAT 1309  
QY  
841 AATTACGAGTTGAAGAGTCTTAACTGACGCTTCTTCTGCGGCGGAGCGCATGAC 900  
Db  
1310 AATTACGAGTTGAAGAGTCTTAACTGACGCTTCTTCTGCGGCGGAGCGCATGAC 1369  
QY  
901 GTAACAGCGGCAATGCGCAACCTCTTTAGAAATGGAACGCTGATTAATTTGGCTG 960  
Db  
1370 GTAACAGCGGCAATGCGCAACCTCTTTAGAAATGGAACGCTGATTAATTTGGCTG 1429  
QY  
961 TCGTCTTTGAGCGGAGTCTGATTAATCAGACAAACCAAGTGGCGGCTTAAACAAACA 1020  
Db  
1430 TCGTCTTTGAGCGGAGTCTGATTAATCAGACAAACCAAGTGGCGGCTTAAACAAACA 1489  
QY  
1021 GGCTCGTTCCGA---CGGATTAATCCAGTGGACGCGCACTATAATCAGAAAGATTG 1077  
Db  
1490 GGCTCGTTCCGA---CGGATTAATCCAGTGGACGCGCACTATAATCAGAAAGATTG 1549  
QY  
1078 GAGATATATACACGCGAGTGAACACCGGATTAATTAATTTGCTGATGAC 1137  
Db  
1550 GATGAATATACACGCGAGTGAACACCGGATTAATTAATTTGCTGATGAC 1609  
QY  
1138 AGCCAAACCGTTGCAACT---GGGCGGCGCAACATCGCTGCTGCTTAAACCTTTCCCGAGT 1194  
Db  
1610 AGCCAAACCGTTGCAACT---GGGCGGCGGCGCAACATCGCTGCTGCTTAAACCTTTCCCGAGT 1669  
QY  
1195 CGGCGTGAAGTTGAAACCTTAAACCGCGAGATTAATTAATTTGCTGATGAC 1254  
Db  
1670 CGGCGTGAAGTTGAAACCTTAAACCGCGAGATTAATTAATTTGCTGATGAC 1729  
QY  
1255 ACTACAGCTCGATTAACACCGGCGGAGATTAATTAATTTGCTGATGAC 1314  
Db  
1730 ACCACAGCAGTATCCAGCATCCGCTGAAACCAACCACTACGCTTCTCACTGTGAC 1789  
QY  
1315 CAAATCCATGAAACGAGCTGTTTACAGCGCGTGAAGATTAATTAATTTGCTGATGAC 1374  
Db  
1790 CAAATCCATGAAACGAGCTGTTTACAGCGCGTGAAGATTAATTAATTTGCTGATGAC 1849  
QY  
1375 ATGACGCTTCAGGAATTTGAATCCGAGTGTCTATGCTTGTGACAAACACCGGCTGACGC 1434

1850 ATGACGCTTCAGGAATTCGAATGCGGAGTGTCTATGCTTGTGACAAACACCGCTTCAGGC 1909  
QY  
1435 AATCTTATAAAGGCTGAGCGGATTTGTTCGGTTTGGCGGCGCAACTGAATCAGGCTTGG 1494  
Db  
1910 AACACTTATAAAGGCTGAGCGGATTTGTTCGGTTTGGCGGCGCAACTGAATCAGGCTTGG 1969  
QY  
1495 CATGTCGGTTACGACATTAATTCGGGCTACCGTGTCCCAATGCGTCCGAAAGTGTATTC 1554  
Db  
1970 CGTTCGGTTACGACATTAATTCGGGCTACCGTGTCCCAATGCGTCCGAAAGTGTATTC 2029  
QY  
1555 ACTTACAAACCGGTTTCGGGTAATTTGGTGTCCCAATCCAACTGAAAGCGGAGCGAGC 1614  
Db  
2030 ACTTACAAACCGGTTTCGGGTAATTTGGTGTCCCAATCCAACTGAAAGCGGAGCGAGC 2089  
QY  
1615 ACCACCCACACCTGCTCTCAGAGGCGGAGCGGAAAGGTAATTTGGTGTCCCACTG 1674  
Db  
2090 ACCACCCACACCTGCTCTCAGAGGCGGAGCGGAAAGGTAATTTGGTGTCCCACTG 2149  
QY  
1675 TATCAAAAATTAACCGCAACTTTCTGTCTGAGAGCAGAGCTGACCAACGAGCGGAT 1734  
Db  
2150 TATCAACCAATTAACCGCAATTTCTGTCTGAGAGCAGAGCTGACCAACGAGCGGAT 2209  
QY  
1735 GTCGGCTGTACTCAGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1794  
Db  
2210 GTCAGCTGTACTCAGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2269  
QY  
1795 CCGGATGCGAGTGAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1854  
Db  
2270 CTGGAATGCGAGTGAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2329  
QY  
1855 CGTCTGAATGTGACAAAAGTAGCGCTTTTCTGCTGAGGCTGGAATTTGTCGGCTG 1914  
Db  
2330 CGTCTGAATGTGACAAAAGTAGCGCTTTTCTGCTGAGGCTGGAATTTGTCGGCTG 2389  
QY  
1915 CTGCGTTATGCGAAAAGCAAACTGTCTGCGGCGCAACACGCTGTCTCCACACGCGCG 1974  
Db  
2390 CTGCGTTATGCGAAAAGCAAACTGTCTGCGGCGCAACACGCTGTCTCCACACGCGCG 2449  
QY  
1975 AAGTGTATGCGGCTGTCTGCTGAGGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 2034  
Db  
2450 AAGTGTATGCGGCTGTCTGCTGAGGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 2509  
QY  
2035 CTGACTTATCTGCGGTGCGAAAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2094  
Db  
2510 CTGACTTATCTGCGGTGCGAAAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2569  
QY  
2095 GCGCGGCTGAGCTTTGCGAGAAAAGGTAATTAATTAATTAATTAATTAATTAATTAATTA 2154  
Db  
2570 GCGTGGGCTGAGCTTTGCGAGAAAAGGTAATTAATTAATTAATTAATTAATTAATTAAT 2629  
QY  
2155 TATGTTTGTGATGATGCTGCGCTTCTACAACTGCGTAAACCTGACTTTGCGTGCAGGC 2214  
Db  
2630 TATGTTTGTGATGATGCTGCGCTTCTACAACTGCGTAAACCTGACTTTGCGTGCAGGC 2689  
QY  
2215 GTATATATGTTTCAACCGCAAAATACACCTTGGGATTTCCCTGCGGCTTTGATAGC 2274  
Db  
2690 GTATATATGTTTCAACCGCAAAATACACCTTGGGATTTCCCTGCGGCTTTGATAGC 2749  
QY  
2275 TACAGCACCAACGCGGCTGAGCCGAGATGGAAGGCTTAGACCGCTACCGGCTCA 2334  
Db  
2750 TACAGCACCAACGCGGCTGAGCCGAGATGGAAGGCTTAGACCGCTACCGGCTCA 2809  
QY  
2335 GCGCGTATTAACGCGTATCGTGGATTTGGAAGTTTGA 2373  
Db  
2810 AGCGGTAATTAACGCGTATCGTGGATTTGGAAGTTTGA 2848

## RESULT 6

US-08-537-361E-1

; Sequence 1, Application US/08537361E

; Patent No. 6121037

; GENERAL INFORMATION:

QY 1315 CAAATCCAAATGGAACGACGTGTTTCAGCAGCGCTGCAGATATCCGTTACGATCATACAAA 1374  
Db 1790 CAAATTCATATGGAACGACGTGTTTCAGTAGCGCGCAGGTATCCGTTACGATCATACAAA 1849  
QY 1375 ATGACGCTCAGGAATGAATCGCGAGTGTATGCTTTGTGACAAAACACCGCTCGAGCC 1434  
Db 1850 ATGACGCTCAGGAATGAATCGCGAGTGTATGCTTTGTGACAAAACACCGCTCGAGCC 1909  
QY 1435 AATACITTAATAAGCTGGAGCGGATTTGTCGGTTTGGCGGGGCACTGAATCAGGCTTGG 1494  
Db 1910 AACACTTAATAAGCTGGAGCGGATTTGTCGGTTTGGCGGGGCACTGAATCAGGCTTGG 1969  
QY 1495 CATGTCGGTTACGACATTTACTTCGGGTACCGTGTCCCAATGCTCCGAGGTGATTTTC 1554  
Db 1970 CGTGTGCGTTACGACATTTACTTCGGGTACCGTGTCCCAATGCTCCGAGGTGATTTTC 2029  
QY 1555 ACTTACAAACACGCTTCGGGTAAATGCTGCTCCCAATCCCAACCTGAAAGCGGAGC 1614  
Db 2030 ACTTACAAACACGCTTCGGGTAAATGCTGCTCCCAATCCCAACCTGAAAGCGGAGC 2089  
QY 1615 ACCACCCACACCTGCTCTGCAAGCGCGCAGGAAAAGGTACTTTGGATGCAACCTG 1674  
Db 2090 ACCACCCACACCTGCTCTGCAAGCGCGCAGGAAAAGGTACTTTGGATGCAACCTG 2149  
QY 1675 TATCAAAACAATTAACGCAACTTCTGCTGTAAGAGCAGAGCTGACCCACGCGCGAT 1734  
Db 2150 TATCAAAACAATTAACGCAACTTCTGCTGTAAGAGCAGAGCTGACCCACGCGCGAT 2209  
QY 1735 GTGCGCTGTACTCAGATGAATTACTACTACGGTATGTGTAGCAATCCTTTATCCGAAAA 1794  
Db 2210 GTGCGCTGTACTCAGATGAATTACTACTACGGTATGTGTAGCAATCCTTTATCCGAAAA 2269  
QY 1795 CCGAATGGCAGATGCAAAATATGATATAAGCCCGAATCCGTTGCTGTGAGCTGACAGCG 1854  
Db 2270 CTGGAATGGCAGATGCAAAATATGATATAAGCCCGAATCCGTTGCTGTGAGCTGACAGCG 2329  
QY 1855 CGTCTGAATGTGCAAAAGTAGGCTTTTCTGCTGAGGCTGGAATTTGCTCGGCTCG 1914  
Db 2330 CGTCTGAATGTGCAAAAGTAGGCTTTTCTGCTGAGGCTGGAATTTGCTCGGCTCG 2389  
QY 1915 CTGGGTTATCGAAAAGCAAACTGTGCGGCGCAACAGCCTGCTGTCCACACAGCGCGCG 1974  
Db 2390 CTGGGTTATCGAAAAGCAAACTGTGCGGCGCAACAGCCTGCTGTCCACCCAGCGCTG 2449  
QY 1975 AAAGTGAATTCGCGTGTGACTAGCAAGCCGAGCGGAAAATGCGGTGTGTTCTCCCGC 2034  
Db 2450 AAAGTGAATTCGCGTGTGACTAGCAAGCCGAGCGGAAAATGCGGTGTGTTCTCCCGC 2509  
QY 2035 CTGACTTTATCTGGTGCAGAAAAGGCCAAGAGCGCAATACACCGTTTATGAAAACAAG 2094  
Db 2510 CTGACTTTATCTGGTGCAGAAAAGGCCAAGAGCGCAATACACCGTTTATGAAAACAAG 2569  
QY 2095 GGCCTGGGTACGCTTTTGCAAGAAAAGGTAAAAGATTACCGCTGGCTGAAACAAGTCGGCT 2154  
Db 2570 GGCCTGGGTACGCTTTTGCAAGAAAAGGTAAAAGATTACCGCTGGCTGAAACAAGTCGGCT 2629  
QY 2155 TATGTGTTTGAATATGATGACGGCTTCTACAACTGGCTTAAACCTGACTTTGGTGCAGGC 2214  
Db 2630 TATGTGTTTGAATATGATGACGGCTTCTACAACTGGCTTAAACCTGACTTTGGTGCAGGC 2689  
QY 2215 GTATATAATGTGTTCAACCGCAAAATACACCACTGGGATTTCCCTGCGCGGTGTTGATAGC 2274  
Db 2690 GTATATAATGTGTTCAACCGCAAAATACACCACTGGGATTTCCCTGCGCGGTGTTGATAGC 2749  
QY 2275 TACAGCACCAACCGCGTGCAGCGAGATGGCAAGGCTTAGACCGCTACCGCGCTCA 2334  
Db 2750 TACAGCACCAACCGCGTGCAGCGAGATGGCAAGGCTTAGACCGCTACCGCGCTCA 2809  
QY 2335 GCGGTTAATTAACCGCTGATCGTGGATTTGGA 2373  
Db 2810 AGCGGTTAATTAACCGCTGATCGTGGATTTGGA 2848

RESULT 5  
US-08-326-670A-1  
; Sequence 1, Application US/08326670A  
; Patent No. 5698438  
; GENERAL INFORMATION:  
; APPLICANT: Stoiljkovic, Igor  
; APPLICANT: So, Magdalene  
; APPLICANT: Hwe, Vivian  
; APPLICANT: Heffron, Fred  
; APPLICANT: Nassif, Xavier  
; TITLE OF INVENTION: A No. 5698438el Bacterial Hemoglobin Receptor Gene and Uses  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Allegretti & Witcoff, Ltd.  
; STREET: 10 South Wacker Drive, Suite 3000  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/326,670A  
; FILING DATE: 18 OCT 1994  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5698438nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 94,784  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-715-1000  
; TELEFAX: 312-715-1234  
; TELEX: 810-221-8317  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3318 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 470..2848  
; US-08-326-670A-1

Query Match 83.6%; Score 1988.6; DB 1; Length 3318;  
Best Local Similarity 90.3%; Pred. No. 0;  
Matches 2149; Conservative 0; Mismatches 224; Indels 6; Gaps 2;  
QY 1 ATGAAACCAATTACATGCTTCTTATGCGCGCTGCTGCGCAGTATTTTCGCAATCCG 60  
Db 470 ATGAAACCAATTACAAATGCTCTCTATCGCGCGCTGCTGCGCAGTATTTTCGCAATCCG 529  
QY 61 GTCTTGGCAGCGATGAGCTGCAACCGAACCACACCCGTTAAAGCAGAGATAAAGAA 120  
Db 530 GTCTTGGCAGCGATGAGCTGCAACCGAACCACACCCGTTAAAGCAGAGATAAAGAA 589  
QY 121 GTCGCGGTTAAAGCAGCGCTTAAATGCGCTTGAACCGTGAACCGTGTCAACCTCGGCGC 180  
Db 590 GTCGCGGTTAAAGCAGCGCTTAAATGCGCTTGAACCGTGTGAACCGTGTCAACCT 649  
QY 181 ATTCAACAGAAATGATACGCAACAAGAGCTTGGTGGTGTACTTCCACCGAGTCCG 240  
Db 650 ATCAACAAAGAAATGATACGCAACAAGAGCTTGGTGGTGTACTTCCACCGAGTCCG 709  
QY 241 TTGAGCGATAGCGCGCCCATCAAAAGGCTTTGCTGTGCGCGCGCTGGAAGCAACCGT 300  
Db 710 TTGAGCGATAGCGCGCCCATCAAAAGGCTTTGCTGTGCGCGCGCTGGAAGCAACCGT 769

**Qy** 2335 GGCCGTAATTACGCCGTATCGCTGGATTGGAAGTTTTCA 2373  
|||||  
**Db** 2810 AGCCGTAATTACGCCGTATCGCTGGAAATGGAAGTTTTAA 2848

## RESULT 4

US-08-817-707-1  
 ; Sequence 1, Application US/08817707  
 ; Patent No. 6277382  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stojiljkovic, Igor  
 ; APPLICANT: So, Magdalene  
 ; APPLICANT: Hwa, Vivian  
 ; APPLICANT: Heffron, Fred  
 ; APPLICANT: Nassif, Xavier  
 ; TITLE OF INVENTION: No. 6277382el Bacterial Hemoglobin Receptor  
 ; TITLE OF INVENTION: Genes and Uses  
 ; NUMBER OF SEQUENCES: 15  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive, 32nd Floor  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/817,707  
 ; FILING DATE: 19-AUG-1997

| Query Match                | 83.7% | Score 1990.2 | DB 3           | Length 3318 |
|----------------------------|-------|--------------|----------------|-------------|
| Best Local Similarity      | 90.4% | Pred. No. 0  |                |             |
| Matches 2150: Conservative |       | 0            | Mismatches 223 | Indels 6    |
| Gaps                       |       |              |                |             |

1 ATGAAACCATTAACATGCTTCTTATTGCGCGCTGTGCGCAGTATTTTCGCAATCCG 60  
470 ATGAAACCATTAACATGCTTCTTATTGCGCGCTGTGCGCAGTATTTTCGCAATCCG 529  
61 GTCTTGGCAGCGGATGAAGCTGCAACCGAAACACACCCCGTTAAGCAGAGATAAAGAA 120  
530 GTCTTTCGCGCAGATGAAGCTGCAACTGAACACACCCCGTTAAGGCAGAGGTAAAGA 589  
121 GTGCGGCTTAAAGACACGCTTAATGCGCGCTGCAACCGTGAACCGTGAACCTCGCGCGC 180  
590 GTGCGGCTTAAAGCCAGCGCAATGCGCGCTGCGCTGTGGAACGGCTCAACCTTAACCT 649  
181 ATTCAACAGGAAATGATACGGCACAACAAAGACTTGGTGGTTACTCCACGACGTCGCG 240

|    |      |  |      |
|----|------|--|------|
| Db | 650  | ATCAAACAAGAAATGATACCGGACAACAAGACTTGGTGGCGTATTTCACCGGATGTCGGC     | 709  |
| Qy | 241  | TTGAGCGATAGCGCGCGCCATCAAAAAGGCTTTGCTGTGCGGCGGTGGAAGCAACCGT       | 300  |
| Db | 710  | TTGAGCGACAGCGCGCCATCAAAAAGGCTTTGCTGTGCGGCGGTGGAAGCAACCGT         | 769  |
| Qy | 301  | GTGCGGTGACGATTTGACGGCGTGAGCGTGCCTGTATTCGAGAGAAACTCATCTGTATGCA    | 360  |
| Db | 770  | GTGCGGTGAGCATAGACGGCGTAAACCTGCCTGATTCGAAAGAAAACCTCGCTGTACGCC     | 829  |
| Qy | 361  | CGTTATGGCAACTTCAACAGCTCGCGCCTGTCTATCGACCCCGAACTCGTGGCGCAACATC    | 420  |
| Db | 830  | CGTTATGGCAACTTCAACAGCTCGCTGTCTATCGACCCCGAACTCGTGGCGCAACATC       | 889  |
| Qy | 421  | GAATCGCGAAGGCGCTGACTCTTTCAATACCGGTAGCGGCGCATTTGGGTGGCGCGT        | 480  |
| Db | 890  | GACATCGTAAAGGGCGGACTCTTTCAATACCGGAGCGCGCCTTTGGGCGGCGGTGTG        | 949  |
| Qy | 481  | AATTACCAACCTGCAAGGACATGATTTGCTGTGTTGGACGACAGCAATTCGCGCTGATG      | 540  |
| Db | 950  | AATTACCAACCTGCAAGGACATGATTTGCTGTGTTGGCTGAAACGGCAGTTTCGCGGTGATG   | 1009 |
| Qy | 541  | ATGAAAAACGGTTACAGCAGCGCAACCGCGAATGGAACAAATACACTTCGGTTTCGGTGTG    | 600  |
| Db | 1010 | ATGAAAAACGGTTACAGCAGCGTAACCGTGAATGGAACAAATACACTTCGGTTTCGGCGTG    | 1069 |
| Qy | 601  | AGCAACGACCGGTGGATGCGCTTTGCTGTATTCGAAACGTCGCGGTCAAGAGCCGAA        | 660  |
| Db | 1070 | AGCAACGACCGGTGGATGCGCTTTGCTGTATTCGAAACGCGCGGCCATGAAACTGAA        | 1129 |
| Qy | 661  | AGCGCGGCGAGCGTGGCTATCCGTTAGAGGGTGTGCGAGCGGAGCAATTATCCGTGCT       | 720  |
| Db | 1130 | AGCGCGGCGAAGCGTGGTATCCGTTAGAGGGTGTGTTAGCGGAGCGAATATCCGTGCT       | 1189 |
| Qy | 721  | TCGTCACGCGGTATCCCTGATCCGTCCAAACCAAAATACCAACTTCTTTGGTAAAGATT      | 780  |
| Db | 1190 | TCTGCGCGGTATTCTGTATCCGTCCCAACAACAATACCAAGCTTCTTTGGTAAAGATT       | 1249 |
| Qy | 781  | GCTTTATCAATCAACGACAGCAACCGCATCGGCCCATTCGTTTAAACGCGCAGCAGGGGCAT   | 840  |
| Db | 1250 | GCTTTATCAATCAACGACCAACCAACCGCATCGGCCCATTCGTTTAAACGCGCAGCAGGGGCAT | 1309 |
| Qy | 841  | AATTACACGATTCGACAGTCTTATACCTGACGCGTCTTCTCTGGCGGCAAGCCGATGAC      | 900  |
| Db | 1310 | AATTACACGATTCGACAGTCTTATACCTGCTTCTTATTTGGCGTGAAGTGAAGTAT         | 1369 |
| Qy | 901  | GTAAACAGACGGCGCAATGCCACCTCTTTTACGAATGGACGCGCTGATTCAAATTTGGCTG    | 960  |
| Db | 1370 | GTCAACAGACGGCGTAAACCAACCTCTTTACGAATGGACGCGCAATCCGACCGGTTG        | 1429 |
| Qy | 961  | TCGCTTTTGAAGCGGACTTCGATATACAGACAACAAAGTGGCGCGGTAAACAAACA         | 1020 |
| Db | 1430 | TCTATGTTAAAGCGGATGTGATATCAAAAAACCAAAAGTATCTGCGGTCAACTACAA        | 1489 |
| Qy | 1021 | GGCTCGTTCGGA - - - CGGATATTCCACTGAGCGCAACTATATATCAGAAAGGATTTG    | 1077 |
| Db | 1490 | GGTTCGTTCCGATGAGGATTTCTTCCACTTGACGCTAACATCAAAAGGACTTTG           | 1549 |
| Qy | 1078 | GAGAAATATACAAACCGCAGCATGGACACCGGATTCAAAACGTTTACTTTTGGGTATGGAC    | 1137 |
| Db | 1550 | GATGAAATCTCAACCGCAGTATGGATACCGCTTCAACCGCATTACCCTGCTTTGGAC        | 1609 |
| Qy | 1138 | AGCACAACGGTTGAACT - - - GGGCGGCAACATCGCTTGTGCTTTAAACTTTTCGCCAGT  | 1194 |
| Db | 1610 | AGGCATCCGTTGCAACTCGGGGGGGGGGACACCGCTGTCTGTTTAAAACTTTTCGCCAGC     | 1669 |
| Qy | 1195 | CGCGGTGAGTTTCAAAACTTAAACCGCAGCATTATTACTTTCAGCGAAAGAGTATCCCGT     | 1254 |
| Db | 1670 | CGCGGTGATTTGAAAACCTTAAACCGCAGCATTATTACTTTCAGCGCGCGTGTGTTGCA      | 1729 |
| Qy | 1255 | ACTACGAGCTCGATTCAACACCCCGTGAAAAACCACTAATTATGGTTTCTCATGTCGAT      | 1314 |
| Db | 1730 | ACCACGAGCATTCAGCATCCGGTGAAAAACCAACCACTACGTTTCTCATGTCGAT          | 1789 |



Db 1141 CAAACCGTTGCAACTGGGCGCCCAACATCGTCTGTGCTTAAACCTTTCGCCAGTCCGGCT 1200  
QY 1201 GAGTTTGAACCTTAAACCGCGACGATTAATCTTTCAGCGAAGAGTATCCCGTACTACC 1260  
Db 1201 GAGTTTGAACCTTAAACCGCGACGATTAATCTTTCAGCGAAGAGTATCCCGTACTACC 1260  
QY 1261 AGCTCGATTCAACACCCCGTGAACCACTAATATATGTTTCTCACTGCTCTCATCAAAATC 1320  
Db 1261 AGCTCGATTCAACACCCCGTGAACCACTAATATGTTTCTCACTGCTCTCATCAATC 1320  
QY 1321 CAATGGAACGAGTGTTCAGACGCGTGCAGATATCCGTTACGATCATACCAAAATGAGC 1380  
Db 1321 CAATGGAACGAGTGTTCAGACGCGTGCAGATATCCGTTACGATCATACCAAAATGAGC 1380  
QY 1381 CCTCAGGAATCAATGCGGAGTGTATGCTTGTGCAAAACACCGCTGCAGCCCAATCT 1440  
Db 1381 CCTCAGGAATCAATGCGGAGTGTATGCTTGTGCAAAACACCGCTGCAGCCCAATCT 1440  
QY 1441 TATAAGGCTGAGCGGATTTGCTGCTTTGGCGGCGCAACTGAATCAGGCTTGGCATGTC 1500  
Db 1441 TATAAGGCTGAGCGGATTTGCTGCTTTGGCGGCGCAACTGAATCAGGCTTGGCATGTC 1500  
QY 1501 GGTACGACATTTCTTCGGCTACCGTGTCCCAATGCTCCGAAGTGTATTTCACTTAC 1560  
Db 1501 GGTACGACATTTCTTCGGCTACCGTGTCCCAATGCTCCGAAGTGTATTTCACTTAC 1560  
QY 1561 AACCAAGTTCGGGTAAATGCTGCTCCCAATCCCACTGAAAGCGGCGAGCACACC 1620  
Db 1561 AACCAAGTTCGGGTAAATGCTGCTCCCAATCCCACTGAAAGCGGCGAGCACACC 1620  
QY 1621 CACACCTGCTCTGCAAGCGCGAGCGAAAGTACTTTGGATGCCAACCTGTATCAA 1680  
Db 1621 CACACCTGCTCTGCAAGCGCGAGCGAAAGTACTTTGGATGCCAACCTGTATCAA 1680  
QY 1681 AACCAATTCGCAACTTTCTGTTCTGAAGAGCAAGAGTGACCAACAGCGGCGATGTCGC 1740  
Db 1681 AGCAATTCGCAAACTTTCTGTTCTGAAGAGCAAGAGTGACCAACAGCGGCGATGTCGC 1740  
QY 1741 TGTACTCAGATCAATTTACTACTACGTTAGTGTGCTGCAATCTTATTCGGAACCGGAA 1800  
Db 1741 TGTACTCAGATCAATTTACTACTACGTTAGTGTGCTGCAATCTTATTCGGAACCGGAA 1800  
QY 1801 TGGCAGATGCAAAATATCGATAAGCGCCGAATCCGTGCTTTGAGCTGACAGCGCGTCTG 1860  
Db 1801 TGGCAGATGCAAAATATCGATAAGCGCCGAATCCGTGCTTTGAGCTGACAGCGCGTCTG 1860  
QY 1861 AATGTGACAAAGTAGCGTCTTTGTTCTGAGGCTGGAATTTGTTGCTGCTGGGT 1920  
Db 1861 AATGTGACAAAGTAGCGTCTTTGTTCTGAGGCTGGAATTTGTTGCTGCTGGGT 1920  
QY 1921 TATGCGAAAGCAAACTGTGCGGCGACACACGCTGCTGTCACACAGCGCGGAAAGTG 1980  
Db 1921 TATGCGAAAGCAAACTGTGCGGCGACACACGCTGCTGTCACACAGCGCGGAAAGTG 1980  
QY 1981 ATTGCGGCTGCTGACTTACGAAAGCGGAGCGGAAATGCGGTGCTTCTCCGCGCTGACT 2040  
Db 1981 ATTGCGGCTGCTGACTTACGAAAGCGGAGCGGAAATGCGGTGCTTCTCCGCGCTGACT 2040  
QY 2041 TATCTGGGTGCAAAAGGCGCAAGACGCGCAATACACCGTTTATGAAACAAAGGCGCG 2100  
Db 2041 TATCTGGGTGCAAAAGGCGCAAGACGCGCAATACACCGTTTATGAAACAAAGGCGCG 2100  
QY 2101 GGTAGCCTTTTCAGAAAGAGTAAAGATTAACCGTGGCTGAAACAGTTCGCTTATG 2160  
Db 2101 GGTAGCCTTTTCAGAAAGAGTAAAGATTAACCGTGGCTGAAACAGTTCGCTTATG 2160  
QY 2161 TTTGATATGTCAGGCTTTCTACAACTGGCTAAACCTGACTTTGCTGCGAGCGGTAT 2220  
Db 2161 TTTGATATGTCAGGCTTTCTACAACTGGCTAAACCTGACTTTGCTGCGAGCGGTAT 2220  
QY 2221 AATGTGTTCAACCGCAAAATACACCACTTGGGATTCCTTGCAGCGGTTGTATAGCTACAGC 2280

Db 2221 AATGTGTTCAACCGCAAAATACACCACTTGGGATTCCTTGCAGCGTTGTATAGCTACACC 2280  
QY 2281 ACCACCAACGCGTCCGAGATGCGCAAGCTTTAGACCGCTACCGGCGCTCAGGCGCT 2340  
Db 2281 ACCACCAACGCGTCCGAGATGCGCAAGCTTTAGACCGCTACCGGCGCTCAGGCGCT 2340  
QY 2341 AATTACGCGGTATCGCTGATTTGGAAGTGGAAATTTGAATTC 2378  
Db 2341 AATTACGCGGTATCGCTGATTTGGAAGTGGAAATTTGAATTC 2378  
  
RESULT 3  
US-08-990-470A-1  
; Sequence 1, Application US/08990470A  
; Patent No. 6123942  
; GENERAL INFORMATION:  
; APPLICANT: Stojiljkovic, Igor  
; APPLICANT: So, Magdalene  
; APPLICANT: Hwa, Vivian  
; APPLICANT: Heffron, Fred  
; APPLICANT: Nassif, Xavier  
; TITLE OF INVENTION: No. 6123942el Bacterial Hemoglobin Receptor  
; TITLE OF INVENTION: Genes and Uses  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boshnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive, 32nd Floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/990,470A  
; FILING DATE: 15-DEC-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 6123942nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 94,784-K  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-0002  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3318 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 470..2845  
; US-08-990-470A-1  
  
Query Match 83.7%; Score 1990.2; DB 3; Length 3318;  
Best Local Similarity 90.4%; Pred. No. 0;  
Matches 2150; Conservative 0; Mismatches 223; Indels 6; Gaps 2;  
  
QY 1 ATGAACCATTTACATGCTTCTTATTCGCGGCTGTCGCGAGTATTTTCGGCAATCCG 60  
Db 470 ATGAACCATTTACAAATGCTCTTATTCGCGGCTGTCGCGAGTATTTTCGGCAATCCG 529  
QY 61 GTCTTGGCAGCGGATGAAGCTGCAACCGCAACCGCTTTAAAGCAGAGATAAAGAA 120  
Db 530 GTCTTGGCAGAGTGAAGCTGCAACTGAACACCGCTTTAAGCAGAGATAAAGCA 589  
QY 121 GTGCGCGTTAAAGACCACTTAAATGCGCTCTTAATGCGCTGCAACCGTGAACCGTCCGCGC 180



QY 2161 TTTGATATGACGGCTTCTCAAACTGGCTAAACCTGACTTTGCGTGCGAGCGTATAT 2220  
DB 2161 TTTGATATGACGGCTTCTCAAACTGGCTAAACCTGACTTTGCGTGCGAGCGTATAT 2220  
QY 2221 AATGTTTCAACCGCAATACACACTTGGATTCCTCGCGCGTGTGTTGTTAGTACAGC 2280  
DB 2221 AATGTTTCAACCGCAATACACACTTGGATTCCTCGCGCGTGTGTTGTTAGTACAGC 2280  
QY 2281 ACCACCAACCGGTGCGACCGAGATGCGAAAGGCTTAGACCGCTACCGCGCTCAGGCCGT 2340  
DB 2281 ACCACCAACCGGTGCGACCGAGATGCGAAAGGCTTAGACCGCTACCGCGCTCAGGCCGT 2340  
QY 2341 AATTACCGCTATCGCTGGATTTGGAATTC 2378  
DB 2341 AATTACCGCTATCGCTGGATTTGGAATTC 2378

RESULT 2  
US-08-537-361E-7  
Sequence 7, Application US/08537361E  
Patent No. 6121037  
GENERAL INFORMATION:  
APPLICANT: Stojiljkovic, Igor  
APPLICANT: So, Magdalene  
APPLICANT: Hwa, Vivian  
APPLICANT: Heffron, Fred  
APPLICANT: Nassif, Xavier  
TITLE OF INVENTION: No. 6121037el Bacterial Hemoglobin Receptor  
TITLE OF INVENTION: Genes and Uses  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
STREET: 300 South Wacker Drive, 32nd Floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/537,361E  
FILING DATE: 02-OCT-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6121037nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 94,784-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2378 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..2370  
US-08-537-361E-7

Query Match 99.2%; Score 2358.8; DB 3; Length 2378;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 2366; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 1 ATGAACCAATACACATGCTTCTTATTCGCGCGCTGCGCAGTATTTTCGGCAATCCG 60  
DB 1 ATGAACCAATACACATGCTTCTTATTCGCGCGCTGCGCAGTATTTTCGGCAATCCG 60

QY 61 GTCTTGGCAGCGGATGAAGCTGCAACCGAAACCAACCCGTTAAAGCAGAGATAAAAGAA 120  
DB 61 GTCTTGGCAGCGGATGAAGCTGCAACCGAAACCAACCCGTTAAAGCAGAGATAAAAGAA 120  
QY 121 GTGCGGTTAAAGACGAGCTTAATGGCTTGCAACCGTGAACGTTCAACCTCGGCGCG 180  
DB 121 GTGCGGTTAAAGACGAGCTTAATGGCTTGCAACCGTGAACGTTCAACCTCGGCGCG 180  
QY 181 ATTCAACAGGAATATGATACGCGACAACAAAGACTTGGTGGTTACTTCCACGAGCTCGGC 240  
DB 181 ATTCAACAGGAATATGATACGCGACAACAAAGACTTGGTGGTTACTTCCACGAGCTCGGC 240  
QY 241 TTGAGCGATAGCGCGCCCATCAAAAGGCTTTGCTGTGCGCGCGTGAAGGCAACCGT 300  
DB 241 TTGAGCGATAGCGCGCCCATCAAAAGGCTTTGCTGTGCGCGCGTGAAGGCAACCGT 300  
QY 301 GTCGTTGTACGATTTGACGCGGTGAGCTGCTGATTCGGAAGAAACTCACTGTATGCA 360  
DB 301 GTCGTTGTACGATTTGACGCGGTGAGCTGCTGATTCGGAAGAAACTCACTGTATGCA 360  
QY 361 CGTTATGGCAACTTCAACAGCTCGCGCTGCTATCGACCCCGAATCTGTCGCGCAATC 420  
DB 361 CGTTATGGCAACTTCAACAGCTCGCGCTGCTATCGACCCCGAATCTGTCGCGCAATC 420  
QY 421 GAAATCGCGAAGGCGCTGACTCTTTCAATACCGGTAGCGCGCATTTGGGTGCGCGGTG 480  
DB 421 GAAATCGCGAAGGCGCTGACTCTTTCAATACCGGTAGCGCGCATTTGGGTGCGCGGTG 480  
QY 481 AATTACCAACCTTCAACAGCATGATTTGCTGTGAGCAGCAGGCAATTTGCGGTTGATG 540  
DB 481 AATTACCAACCTTCAACAGCATGATTTGCTGTGAGCAGCAGGCAATTTGCGGTTGATG 540  
QY 541 ATGAAAAACGTTTACAGCAGCGCAACCGCAATGACAAATACACTCGGTTTCGGTGTG 600  
DB 541 ATGAAAAACGTTTACAGCAGCGCAACCGCAATGACAAATACACTCGGTTTCGGTGTG 600  
QY 601 AGCAACGACCGCTGGATCGCTTTGCTGTATTCGCAACGTCGCGTCAATGACACGAA 660  
DB 601 AGCAACGACCGCTGGATCGCTTTGCTGTATTCGCAACGTCGCGTCAATGACACGAA 660  
QY 661 AGCGCGGCGAGCGTGCTATCCGTAGAGGCTGCTGGCAGCGGAGCAATATCCGTTG 720  
DB 661 AGCGCGGCGAGCGTGCTATCCGTAGAGGCTGCTGGCAGCGGAGCAATATCCGTTG 720  
QY 721 TCGTCAACGCGGTATCCGTGATCCGTTCCAAACACAAATACCAAACTTCTTTGGTAAAGAT 780  
DB 721 TCGTCAACGCGGTATCCGTGATCCGTTCCAAACACAAATACCAAACTTCTTTGGTAAAGAT 780  
QY 781 GCTTATCAATCAACGACAGCAGCAGCGCATCGGCCCATCGTTTAAACGCGCAGCGGCGAT 840  
DB 781 GCTTATCAATCAACGACAGCAGCAGCGCATCGGCCCATCGTTTAAACGCGCAGCGGCGAT 840  
QY 841 AATTACAGATTGAAGCTTTAATACCTGACCGCTTCTTCTGGCGCGAGCGGATGAC 900  
DB 841 AATTACAGATTGAAGCTTTAATACCTGACCGCTTCTTCTGGCGCGAGCGGATGAC 900  
QY 901 GTAAACAGACGCGCAATGCAACCTTTTACGAATGAGCGCTGATTCAAATTTGCTG 960  
DB 901 GTAAACAGACGCGCAATGCAACCTTTTACGAATGAGCGCTGATTCAAATTTGCTG 960  
QY 961 TCGTCTTTGAAAGCGGATTCGATTTATCAGAACCAACAAAGTGGCGGTTTAAACAAA 1020  
DB 961 TCGTCTTTGAAAGCGGATTCGATTTATCAGAACCAACAAAGTGGCGGTTTAAACAAA 1020  
QY 1021 GGTCTGTTCCGAGCGGATTTTCCACTGACCGCAACTATATCAGAGGATTTGAG 1080  
DB 1021 GGTCTGTTCCGAGCGGATTTTCCACTGACCGCAACTATATCAGAGGATTTGAG 1080  
QY 1081 AATATATACACCGCAGCATGGACACCCGATTCGAAAGCTTTTACTTTTCGCTATGGACAGC 1140  
DB 1081 AATATATACACCGCAGCATGGACACCCGATTCGAAAGCTTTTACTTTTCGCTATGGACAGC 1140  
QY 1141 CAACCGTTGCAACTGGGCGGCCAACATCGCTTGTGCTTTAAACTTTTCGCCAGTCGCGCT 1200

| Matches 2369; Conservative 0; Mismatches 9; Indels 0; Gaps 0; |      |   |      |
|---|------|---|------|
| Qy  | 1    | ATGAAACCAATTACACATGCTTCTCTATTTGCGCGCTGGTCCGCGAGTATTTTCGGCAATCCG | 60   |
| Db  | 1    | ATGAAACCAATTACACATGCTTCTCTATTTGCGCGCTGGTCCGCGAGTATTTTCGGCAATCCG | 60   |
| Qy  | 61   | GTCTTGGCAGCGGATGAAGCTGCAACCGAACCACACCCGGTTAAAGCAGAGATAAAGAA     | 120  |
| Db  | 61   | GTCTTGGCAGCGGATGAAGCTGCAACCGAACCACACCCGGTTAAAGCAGAGATAAAGAA     | 120  |
| Qy  | 121  | GTGCGGTTAAAGACAGCTTAATGCGCTGCAACCGTGGAAAGCTGTCAACCTCGGCGCG      | 180  |
| Db  | 121  | GTGCGGTTAAAGACAGCTTAATGCGCTGCAACCGTGGAAAGCTGTCAACCTCGGCGCG      | 180  |
| Qy  | 181  | ATTCAACAGGAAATGATACGCGACCAACAAAGACTTTGGTGCCTTCTCCACCGACGTCGGC   | 240  |
| Db  | 181  | ATTCAACAGGAAATGATACGCGACCAACAAAGACTTTGGTGCCTTCTCCACCGACGTCGGC   | 240  |
| Qy  | 241  | TTGAGCGATAGCGCGCCCATCAAAAAGCTTTTGTCTGTGCGCGGCTGGAAAGGCAACCGT    | 300  |
| Db  | 241  | TTGAGCGATAGCGCGCCCATCAAAAAGCTTTTGTCTGTGCGCGGCTGGAAAGGCAACCGT    | 300  |
| Qy  | 301  | GTGCGGTGACGATTTGACGGCGTGAAGCTGCTGATTTGGAAAGAACTCACTGTATGCA      | 360  |
| Db  | 301  | GTGCGGTGACGATTTGACGGCGTGAAGCTGCTGATTTGGAAAGAACTCACTGTATGCA      | 360  |
| Qy  | 361  | CGTTATGGCAACTTCAACAGCTCGCGCCCTGTCTATCGACCCCGAACTCGTGGCGCAACATC  | 420  |
| Db  | 361  | CGTTATGGCAACTTCAACAGCTCGCGCCCTGTCTATCGACCCCGAACTCGTGGCGCAACATC  | 420  |
| Qy  | 421  | GAATCGGAAGGGCGCTCACTCTTTCAATACCGGTAGCGCGCATTTGGGTGGCGGGTG       | 480  |
| Db  | 421  | GAATCGGAAGGGCGCTCACTCTTTCAATACCGGTAGCGCGCATTTGGGTGGCGGGTG       | 480  |
| Qy  | 481  | AATTACCAACCTGCAAGGACATGATTTGCTGTTGGACGACGCAATTCGGCGTGTATG       | 540  |
| Db  | 481  | AATTACCAACCTGCAAGGACATGATTTGCTGTTGGACGACGCAATTCGGCGTGTATG       | 540  |
| Qy  | 541  | ATGAAAAACGGTTACAGACGCGCAACCGCAATGGACAAATACACTCGGTTTCGGTGTG      | 600  |
| Db  | 541  | ATGAAAAACGGTTACAGACGCGCAACCGCAATGGACAAATACACTCGGTTTCGGTGTG      | 600  |
| Qy  | 601  | AGCAACGACCGGTGATCGCGTTTGTGTATTCGCAACGTCGCGGTCAATGAGACCGAA       | 660  |
| Db  | 601  | AGCAACGACCGGTGATCGCGTTTGTGTATTCGCAACGTCGCGGTCAATGAGACCGAA       | 660  |
| Qy  | 661  | AGCGCGGCGAGGTGGCTATCCGGTAGAGGTGTCTGGCAGCGGAGCAATTCCTGGGT        | 720  |
| Db  | 661  | AGCGCGGCGAGGTGGCTATCCGGTAGAGGTGTCTGGCAGCGGAGCAATTCCTGGGT        | 720  |
| Qy  | 721  | TCGTACCGGGTATCCCTGATCCGTCCAAACAAATACCAACTCTCTGGGTAAAGTT         | 780  |
| Db  | 721  | TCGTACCGGGTATCCCTGATCCGTCCAAACAAATACCAACTCTCTGGGTAAAGTT         | 780  |
| Qy  | 781  | GCTTATCAAATCAACGACCAAGCACCGCATTCGCCCATCGTTTAAACGGCCAGCGGGCAT    | 840  |
| Db  | 781  | GCTTATCAAATCAACGACCAAGCACCGCATTCGCCCATCGTTTAAACGGCCAGCGGGCAT    | 840  |
| Qy  | 841  | AATTACAGATTGAAGAGCTTATTAACCTGACCGCTTCTTCTGGCGGAGCGCGATGAC       | 900  |
| Db  | 841  | AATTACAGATTGAAGAGCTTATTAACCTGACCGCTTCTTCTGGCGGAGCGCGATGAC       | 900  |
| Qy  | 901  | GTAACACAGCGCGCAATCCCAACTCTTTTACGAATGGAGCCCTGATTCAAATTTGGCTG     | 960  |
| Db  | 901  | GTAACACAGCGCGCAATCCCAACTCTTTTACGAATGGAGCCCTGATTCAAATTTGGCTG     | 960  |
| Qy  | 961  | TCGCTCTTTGAAGCGGACTTTCGATTTATCAGACCAACCAAGTGGCGCGGTTTAAACAAA    | 1020 |
| Db  | 961  | TCGCTCTTTGAAGCGGACTTTCGATTTATCAGACCAACCAAGTGGCGCGGTTTAAACAAA    | 1020 |
| Qy  | 1021 | GGCTCGTTCGCGAGGATTTTCCACTGGACCGCGCAACTATATACAGAGGATTTGGAG       | 1080 |
| Db  | 1021 | GGCTCGTTCGCGAGGATTTTCCACTGGACCGCGCAACTATATACAGAGGATTTGGAG       | 1080 |

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2005, 19:52:52 ; Search time 405 Seconds  
(without alignments)  
9607.570 Million cell updates/sec

Title: US-09-665-358-7

Perfect score: 2378

Sequence: 1 ATGAACCAATTACACATGCT.....GATTGGAAGTTTGAATTC 2378

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*
- 2: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*
- 5: /cgn2\_6/ptodata/1/ina/PCFUS COMB.seq:\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length  | DB ID | Description          |
|------------|--------|-------------|---------|-------|----------------------|
| 1          | 2363.6 | 99.4        | 2378    | 3     | US-08-817-707-7      |
| 2          | 2358.8 | 99.2        | 2378    | 3     | US-08-537-361E-7     |
| 3          | 1990.2 | 83.7        | 3318    | 3     | US-08-990-470A-1     |
| 4          | 1990.2 | 83.7        | 3318    | 3     | US-08-817-707-1      |
| 5          | 1988.6 | 83.6        | 3318    | 1     | US-08-326-670A-1     |
| 6          | 1985.4 | 83.5        | 3319    | 3     | US-08-537-361E-1     |
| 7          | 1983.8 | 83.4        | 2379    | 3     | US-08-537-361E-5     |
| 8          | 1958.4 | 82.4        | 2373    | 3     | US-08-817-707-5      |
| 9          | 1849.6 | 77.8        | 2376    | 3     | US-08-817-707-3      |
| 10         | 1848   | 77.7        | 2376    | 3     | US-08-537-361E-3     |
| 11         | 68.4   | 2.9         | 1401    | 4     | US-09-252-991A-14258 |
| 12         | 68.4   | 2.9         | 1767    | 4     | US-09-252-991A-14416 |
| 13         | 68.4   | 2.9         | 3738    | 4     | US-09-252-991A-14364 |
| 14         | 65.4   | 2.8         | 2809    | 3     | US-08-448-194-5      |
| 15         | 65.4   | 2.8         | 2809    | 3     | US-08-867-921-5      |
| 16         | 65     | 2.7         | 3286    | 2     | US-08-363-124A-1     |
| 17         | 63.8   | 2.7         | 2955    | 2     | US-08-867-941-9      |
| 18         | 63.8   | 2.7         | 2955    | 3     | US-09-074-658-9      |
| 19         | 63.8   | 2.7         | 3000    | 2     | US-08-867-941-8      |
| 20         | 63.8   | 2.7         | 3000    | 3     | US-09-074-658-8      |
| 21         | 63.8   | 2.7         | 7641    | 2     | US-08-867-941-6      |
| 22         | 63.8   | 2.7         | 7641    | 3     | US-09-074-658-6      |
| 23         | 63.6   | 2.7         | 2370    | 4     | US-09-252-991A-10440 |
| 24         | 62.2   | 2.6         | 3012    | 4     | US-09-540-236-837    |
| 25         | 60.8   | 2.6         | 2826    | 3     | US-08-624-655A-1     |
| 26         | 60.8   | 2.6         | 1830121 | 4     | US-09-557-884-1      |
| 27         | 60.8   | 2.6         | 1830121 | 4     | US-09-643-990A-1     |

28 60.6 2.5 49617 4 US-09-596-002-28 Sequence 28, Appli  
29 59.4 2.5 2859 5 PCT-US96-05320A-637 Sequence 637, Appli  
30 59.4 2.5 3252 3 US-08-476-102A-4 Sequence 4, Appli  
31 59.4 2.5 3252 5 PCT-US96-05320A-714 Sequence 714, Appli  
32 57.4 2.4 2955 2 US-08-867-941-4 Sequence 4, Appli  
33 57.4 2.4 2955 3 US-09-074-658-4 Sequence 4, Appli  
34 57.4 2.4 3000 2 US-08-867-941-3 Sequence 3, Appli  
35 57.4 2.4 3000 3 US-09-074-658-3 Sequence 3, Appli  
36 57.4 2.4 7650 2 US-08-867-941-1 Sequence 1, Appli  
37 57.4 2.4 7650 3 US-09-074-658-1 Sequence 1, Appli  
38 55.8 2.3 2696 1 US-07-961-522-1 Sequence 1, Appli  
39 55.8 2.3 2696 1 US-08-217-438-1 Sequence 1, Appli  
40 55.8 2.3 2696 1 US-08-321-978-1 Sequence 1, Appli  
41 55.8 2.3 2696 2 US-08-710-584-1 Sequence 1, Appli  
42 54.8 2.3 2800 3 US-08-448-194-3 Sequence 3, Appli  
43 54.8 2.3 2800 3 US-08-867-921-3 Sequence 3, Appli  
44 54.8 2.3 3537 2 US-08-363-124A-3 Sequence 3, Appli  
45 54 2.3 2397 4 US-09-489-039A-2874 Sequence 2874, Ap

## ALIGNMENTS

RESULT 1  
US-08-817-707-7  
; Sequence 7, Application US/08817707  
; Patent No. 6277382  
; GENERAL INFORMATION:  
; APPLICANT: Stojiljkovic, Igor  
; APPLICANT: So, Magdalena  
; APPLICANT: Hwa, Vivian  
; APPLICANT: Heffron, Fred  
; APPLICANT: Nassif, Xavier  
; TITLE OF INVENTION: No. 6277382el Bacterial Hemoglobin Receptor  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive, 32nd Floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/817,707  
; FILING DATE: 19-AUG-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 6277382nat, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 94,784-J  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-0002  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2378 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..2370  
; US-08-817-707-7

Query Match 99.4%; Score 2363.6; DB 3; Length 2378;  
Best Local Similarity 99.6%; Pred. No. 0;



|           |   |      |
|-----------|---|------|
| DR        | P-PSDB; ABP76975.   |      |
| XX        |   |      |
| PT        | New protein from <i>Neisseria gonorrhoeae</i> , useful for the manufacture of a |      |
| PT        | medicament for treating or preventing <i>N. gonorrhoeae</i> infection.          |      |
| XX        |   |      |
| PS        | Disclosure; Page 223; 815pp; English.   |      |
| XX        |   |      |
| CC        | The present invention relates to proteins from <i>Neisseria gonorrhoeae</i> .   |      |
| CC        | Also disclosed are the nucleic acid molecules encoding the proteins and         |      |
| CC        | antibodies that specifically bind to the proteins. The composition              |      |
| CC        | comprising the protein, nucleic acid or antibody is useful for the              |      |
| CC        | manufacture of a medicament for treating or preventing <i>N. gonorrhoeae</i>    |      |
| CC        | infection, this may be in the form of a vaccine or gene therapy.                |      |
| CC        | Sequences given in records ABZ37706-ABZ42016 represent nucleic acid             |      |
| CC        | molecules of the invention  |      |
| XX        |   |      |
| SQ        | Sequence 348 BP; 93 A; 87 C; 89 G; 79 T; 0 U; 0 Other;                          |      |
|           |   |      |
|           | Query Match 14.5%; Score 344.8; DB 10; Length 348;                              |      |
|           | Best Local Similarity 99.4%; Pred. No. 6.6e-96;                                 |      |
|           | Matches 346; Conservative 0; Mismatches 2; Indels 0; Gaps 0;                    |      |
| Qy        | 2023 GTGTTCTCCCGCCTGACTTATCTGGGTGCGAAAGGCCAAAGACGCGCAATACACCGTT                 | 2082 |
| Db        | 1 GTGTTCTCCCGCCTGACTTATCTGGGTGCGAAAGGCCAAAGACGCGCAATACACCGTT                    | 60   |
| Qy        | 2083 TATGAAACAAAGGCCGCGGTACGCCCTTTCGAGAAAAAGGTAAAGATTACCCGTGGCTG                | 2142 |
| Db        | 61 TATGAAACAAAGGCCGCGGTACGCCCTTTCGAGAAAAAGGTAAAGATTACCCGTGGCTG                  | 120  |
| Qy        | 2143 AACAAAGTCGGCTTATGTGTTTGATATGATGACGGCTTCTACAACTGGCTAAAAACCTGACT             | 2202 |
| Db        | 121 AACAAAGTCGGCTTATGTGTTTGATATGATGACGGCTTCTACAACTGGCTAAAAACCTGACT              | 180  |
| Qy        | 2203 TTGCGTGCAGCGATATAATATGTTTCAACCGGAAATACACCACTTGGGATTCCTCGCG                 | 2262 |
| Db        | 181 TTGCGTGCAGCGATATAATATGTTTCAACCGGAAATACACCACTTGGGATTCCTCGCG                  | 240  |
| Qy        | 2263 GGTTTGATAGCTACAGCACCAACCGCGTGCAGCGAGATGCAAGGCTTAGACCGC                     | 2322 |
| Db        | 241 GGTTTGATAGCTACAGCACCAACCGCGTGCAGCGAGATGCAAGGCTTAGACCGC                      | 300  |
| Qy        | 2323 TACCGCGCCTCAGGCGCGTAATTACGCGGTATCGCTGGATTGGAAGTTT                          | 2370 |
| Db        | 301 TACCGCGCCTCAGGCGCGTAATTACGCGGTATCGCTGGATTGGAAGTTT                           | 348  |
| RESULT 12 |   |      |
| ABZ37965  |   |      |
| ID        | ABZ37965 standard; DNA; 222 BP.   |      |
| XX        |   |      |
| AC        | ABZ37965;   |      |
| XX        |   |      |
| DT        | 07-MAR-2003 (first entry)   |      |
| XX        |   |      |
| DE        | <i>N. gonorrhoeae</i> nucleotide sequence SEQ ID 519.                           |      |
| XX        |   |      |
| KW        | Antibacterial; infection; vaccine; gene therapy; gene; ds.                      |      |
| XX        |   |      |
| OS        | <i>Neisseria gonorrhoeae</i> .  |      |
| XX        |   |      |
| PN        | WO200279243-A2.   |      |
| XX        |   |      |
| FD        | 10-OCT-2002.  |      |
| XX        |   |      |
| XX        | 12-FEB-2002; 2002WO-IB002069.   |      |
| PF        |   |      |
| XX        |   |      |
| PR        | 12-FEB-2001; 2001GB-00003424.   |      |
| XX        |   |      |
| FA        | (CHIR-) CHIRON SPA.   |      |
| XX        |   |      |
| PI        | Fontana MR, Pizza M, Masignani V, Monaci E;                                     |      |
| XX        |   |      |

QY 781 GCTTATCAATCAACGACACGACGATCGGCCATCGTCTTAAAGCGCCAGCGGGCAT 840  
 DB |||||  
 QY 781 GCTTATCAATCAACGACACGACGATCGGCCATCGTCTTAAAGCGCCAGCGGGCAT 840  
 DB |||||  
 QY 841 AATTACACGATT 852  
 DB |||||  
 QY 841 AATTACACGATT 852  
 DB |||||

RESULT 10  
 ABZ37953  
 ID ABZ37953 standard; DNA; 864 BP.  
 XX  
 AC ABZ37953;  
 DT 07-MAR-2003 (first entry)  
 XX  
 DE N. gonorrhoeae nucleotide sequence SEQ ID 495.  
 XX  
 KW Antibacterial; infection; vaccine; gene therapy; gene; ds.  
 XX  
 OS Neisseria gonorrhoeae.  
 XX  
 PN WO200279243-A2.  
 XX  
 PD 10-OCT-2002.  
 XX  
 PF 12-FEB-2002; 2002WO-IB002069.  
 XX  
 PR 12-FEB-2001; 2001GB-00003424.  
 XX  
 PA (CHIR-) CHIRON SPA.  
 XX  
 PI Fontana MR, Pizza M, Massignani V, Monaci E;  
 XX  
 DR WPI; 2003-058415/05.  
 DR P-PSDB; ABP76983.  
 XX  
 PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a  
 PT medicament for treating or preventing N. gonorrhoeae infection.  
 XX  
 PS Disclosure; Page 224; 815pp; English.  
 CC  
 CC The present invention relates to proteins from Neisseria gonorrhoeae.  
 CC Also disclosed are the nucleic acid molecules encoding the proteins and  
 CC antibodies that specifically bind to the proteins. The composition  
 CC comprising the protein, nucleic acid or antibody is useful for the  
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae  
 CC infection, this may be in the form of a vaccine or gene therapy.  
 CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid  
 CC molecules of the invention  
 XX  
 SQ Sequence 864 BP; 228 A; 224 C; 209 G; 203 T; 0 U; 0 Other;  
 Query Match 35.1%; Score 833.8; DB 10; Length 864;  
 Best Local Similarity 98.5%; Pred. No. 1.8e-247;  
 Matches 852; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 1171 TTGTGCTTAAACCTTCCGAGTCGGCTGAGTTGAAACCTTAAACCGGACGATTAT 1230  
 DB 1 TTGTGCTTAAACCTTCCGAGTCGGCTGAGTTGAAACCTTAAACCGGACGATTAT 60  
 QY 1231 TACTTTCAGGGAAGAGTATCCCGTACTACAGCTCGATTTCACACCCCGTGAACCACT 1290  
 DB |||||  
 QY 61 TACTTTCAGGGAAGAGTATCCCGTACTACAGCTCGATTTCACACCCCGTGAACCACT 120  
 DB |||||  
 QY 1291 AATTATGGTTTCTCACTGTCTGATCAAAATCCAAATGGAACGAGTGTTCAGCAGCGGTGCA 1350  
 DB |||||  
 QY 121 AATTATGGTTTCTCACTGTCTGATCAAAATCCAAATGGAACGAGTGTTCAGCAGCGGTGCA 180  
 DB |||||  
 QY 1351 GATATCCGTTAGATCATACCAAAATGAGCGCTCAGGAATTAATGAGCGAGTGTATGCT 1410  
 DB |||||  
 DB 181 GATATCCGTTAGATCATACCAAAATGAGCGCTCAGGAATTAATGAGCGAGTGTATGCT 240

QY 1411 TGTGACAAAACACGCTGCGAGCCAAATCTTATTAAGGCTGGAGCGGATTTCTCGGTTTG 1470  
 DB |||||  
 QY 241 TGTGACAAAACACGCTGCGAGCCAAATCTTATTAAGGCTGGAGCGGATTTCTCGGTTTG 300  
 DB |||||  
 QY 1471 GCGGCGCAACTGAATCAGGCTTGGCATGTCGGTTACGACATTTCTTCCGGCTACCGTGTG 1530  
 DB |||||  
 QY 301 GCGGCGCAACTGAATCAGGCTTGGCATGTCGGTTACGACATTTCTTCCGGCTACCGTGTG 360  
 DB |||||  
 QY 1531 CCCAATGCTCGGAAGTGTATTTCACTTACAACACGCTTCCGGTAAATGGCTGCCCAAT 1590  
 DB |||||  
 QY 361 CCCAATGCTCGGAAGTGTATTTCACTTACAACACGCTTCCGGTAAATGGCTGCCCAAT 420  
 DB |||||  
 QY 1591 CCCAATGCTCGGAAGTGTATTTCACTTACAACACGCTTCCGGTAAATGGCTGCCCAAT 1650  
 DB |||||  
 QY 421 TCCAACCTGAAAGCGGAGCGGAGCACACCCCTGTCTCTGCAAGGCGCGAGCGAA 480  
 DB |||||  
 QY 1651 AAAGTACTTTGGATGCCAATCTGTATCAAAACAATTAACCGCAACTTCTTGTCTGAAGAG 1710  
 DB |||||  
 QY 481 AAAGTACTTTGGATGCCAATCTGTATCAAAACAATTAACCGCAACTTCTTGTCTGAAGAG 540  
 DB |||||  
 QY 1711 CAGAAGCTGACACCGGCGGATGTCGGCTGATCAGATGAATTAATCTACTACCGTATG 1770  
 DB |||||  
 QY 541 CAGAAGCTGACACCGGCGGATGTCGGCTGATCAGATGAATTAATCTACTACCGTATG 600  
 DB |||||  
 QY 1771 TGTAGCAATCTTATTCGAAACACCGGAATGCGAGATGCAAAATATCGATAAGGCCCGA 1830  
 DB |||||  
 QY 601 TGTAGCAATCTTATTCGAAACACCGGAATGCGAGATGCAAAATATCGATAAGGCCCGA 660  
 DB |||||  
 QY 1831 ATCCGTGCTTTGAGCTGACAGGCGCTCTGAAATGTGACAAAGTAGCGTCTTTTGTCTCT 1890  
 DB |||||  
 QY 661 ATCCGTGCTTTGAGCTGACAGGCGCTCTGAAATGTGACAAAGTAGCGTCTTTTGTCTCT 720  
 DB |||||  
 QY 1891 GAGGCTCGAAATTTGTCGGCTCGCTGCTTATGCGAAAGCAAACTGTGCGGCGCACAC 1950  
 DB |||||  
 QY 721 GAGGCTCGAAATTTGTCGGCTCGCTGCTTATGCGAAAGCAAACTGTGCGGCGCACAC 780  
 DB |||||  
 QY 1951 AGCTGCTGTCTCACACAGCGCCGAAAGTGTGCGGCTGCGACTACGAAAGCCCGGAGC 2010  
 DB |||||  
 QY 781 AGCTGCTGTCTCACACAGCGCCGAAAGTGTGCGGCTGCGACTACGAAAGCCCGGAGC 2035  
 DB |||||  
 QY 2011 GAAAAATGGGTGTGTTCTCCGCGC 2035  
 DB |||||  
 DB 840 GAAAAATGGGTGTGTTCTCCGCGC 864

RESULT 11  
 ABZ37945  
 ID ABZ37945 standard; DNA; 348 BP.  
 XX  
 AC ABZ37945;  
 XX  
 DT 07-MAR-2003 (first entry)  
 XX  
 DE N. gonorrhoeae nucleotide sequence SEQ ID 479.  
 XX  
 KW Antibacterial; infection; vaccine; gene therapy; gene; ds.  
 XX  
 OS Neisseria gonorrhoeae.  
 XX  
 PN WO200279243-A2.  
 XX  
 PD 10-OCT-2002.  
 XX  
 PF 12-FEB-2002; 2002WO-IB002069.  
 XX  
 PR 12-FEB-2001; 2001GB-00003424.  
 XX  
 PA (CHIR-) CHIRON SPA.  
 XX  
 PI Fontana MR, Pizza M, Massignani V, Monaci E;  
 XX  
 DR WPI; 2003-058415/05.





(UYOR-) UNIV OREGON HEALTH SCI.

XX Stojiljkovic I, So M, Hwa V, Heffron F, Nassif X;  
 XX WPI; 1996-222006/22.  
 DR P-PSDB; AAR95566.

XX DNA encoding Neisseria haemoglobin receptor proteins - for use in  
 PT preparing polypeptide(s) and antibodies for vaccines for e.g. meningitis.  
 XX Claim 3; Page 52-55; 104pp; English.

XX A DNA sequence (AAT26998) codes for the bacterial haemoglobin receptor  
 CC (Hmbr) (AAR95566) of N. meningitidis serotype A. It was obt'd. by PCR  
 CC amplification of genomic DNA using primers (AAT27001-02) based on the N.  
 CC meningitidis serotype C hmbR gene (see also AAT26997). The gene can be  
 CC used to prepare recombinant haemoglobin receptor protein, useful as a  
 CC vaccine for meningitis. It can be expressed in attenuated Salmonella host  
 CC cells. It can also be utilised as a probe to detect infection in humans.  
 CC hmbR genes for N. meningitidis serotype B (AAT26999) and for N.  
 CC gonorrhoeae (AAT27000) were similarly obt'd. (Updated on 25-MAR-2003 to  
 CC correct PR field.) (Updated on 25-MAR-2003 to correct PI field.) (Updated  
 XX on 27-AUG-2003 to correct OS field.)

XX SQ Sequence 2376 BP; 632 A; 622 C; 621 G; 501 T; 0 U; 0 Other;

Query Match 77.6%; Score 1846.4; DB 2; Length 2376;  
 Best Local Similarity 86.4%; Pred. No. 0;  
 Matches 2052; Conservative 0; Mismatches 321; Indels 3; Gaps 1

QY 1 ATGAAACCAATTACACATGCTTCTATTGCGCGCTGCTGCGCAGTATTTTCGGCAATCCG 60  
 DB |||||  
 QY 1 ATGAAACCAATTACAAATGCCCTATCGCGCGCTGCTGCGCAGTATTTTCGGCAATCCG 60  
 DB |||||

QY 61 GTCTTGGCAGCGGATGAAGCTGCACCGAAACACACCCGTTAAACGACGAGATAAAGAA 120  
 DB |||||

QY 61 GTCTTGGCGGAGATGAAGCTGCACCTGAAACACACCCGTTAAGCGAGAGTAAAGCA 120  
 DB |||||

QY 121 GTGCGCGTTAAAGACACAGCTTAAATGCGCGCTGCACACCGTGAAACGTTCAACCTCGGCGCG 180  
 DB |||||

QY 121 GTGCGGTTAAAGGTCAGCGCAATGGCGCTGCGGCTGTGGAACGCGTCAACCTTAACCGT 180  
 DB |||||

QY 181 ATTCACAGGAAATGATACGCGACAAAGAGCTTGTGCGTTACTCCACCGACGTCGCG 240  
 DB |||||

QY 181 ATCAAAACAGAAATGATACGCGACAAATAAGACTTGTGCGCTATTCCACCGATGTCGCG 240  
 DB |||||

QY 241 TTGACGATAGCGCGCGCCATCAAAAGGCTTTCGTGCGCGCGCTGGAAGCAACCGT 300  
 DB |||||

QY 241 TTGAGCGACAGAGCGCGTCATCAAAAAGGCTTTGCCATTCGCGCGCGTGAAGCGACCGT 300  
 DB |||||

QY 301 GTGCGGTGCAGCATTGACGGGTGAGCTGCGCTGATTTCGGAAGAAAACTCAGTGTATGCA 360  
 DB |||||

QY 301 GTGCGGTGATGATTGACGGGTGAACCTGCTGATTCGGAAGAAAACTGCTGTAGCGC 360  
 DB |||||

QY 361 CGTTATGGCAACTTCAACAGCTCGCGCGCTGTCTATGACCCCGAACTCGTGGCGCAACATC 420  
 DB |||||

QY 361 CGTTATGGCAACTTCAACAGCTCGCGCTGTCTATGACCCCGAACTCGTGGCGCAACATC 420  
 DB |||||

QY 421 GAAATCGCGAGGGCGCTGACTCTTTCAATACCGGTAGCGCGGATTTGGGTGGCGCGGTG 480  
 DB |||||

QY 421 GACATCGTAAAGGGGCGGACTCTTTCAATACCGGAGCGCGGCTTTGGGCGCGCGGTG 480  
 DB |||||

QY 481 AATTACCAAAACCTCGCAAGGACATGATTTGCTGTGTGAACGACAGGCAATTCGCGCGTGTG 540  
 DB |||||

QY 481 AATTACCAAAACCTCGCAAGGACATGATTTGCTGTGTGAACGACAGGCAATTCGCGCGTGTG 540  
 DB |||||

QY 541 ATGAAAAACGGTTACAGCAGCGCAACCGCGAATGGAATAATACACTCGGTTTCGGTGTG 600  
 DB |||||

QY 541 ATGAAAAACGGTTACAGCAGCGTAAACCGTGAATGGAATAATACCTCCGTTTCGCGGTG 600  
 DB |||||

QY 601 AGCAACGACCGGTGGATGCGCGCTTTTGTGTATTTTCGCAACGTCGCGGTATGAGACGAA 660  
 DB |||||

QY 601 AGCAACGACCGGTGGATGCGCGCTTTTGTGTATTTTCGCAACGCGCGCGCATGAAATGAA 660  
 DB |||||

```
Db 601 AGTAACAGCGCGTGGATGCTGCTTTGCTGATTCGCAACGGCGCGCCATGAACCGAA 660
Qy 661 AGCGCGGCGAGCGTGGCTATCCGGTAGAGGGTCTCGCAGCGAGCAATATCCGTGGT 720
Db 661 AGCGCGGCGCAACCGCGGTATCCGGTAGAGGGTCCGGTAAAGAAACGAATATCCGCGGT 720
Qy 721 TCGTACAGCGGTATCCCGTATCCGTGCAAAACACAAATACCAAACTCTCTGGGTAAAGATT 780
Db 721 TCCGCGCGGCGATCCCGGATCCGTGCAAAACCAAAATACCAAACTCTCTGGGTAAAGATT 780
Qy 781 GCTTATCAAAATCAACGACGACCGCATCGGCCATCGTTTAAACGCGCAGCGAGGCGCAT 840
Db 781 GCTTATCAAAATCAACGACGACCAACCGCATCGGCCATCGCTCAACGGTTCAGCAGGCGCAT 840
Qy 841 AATTACAGATTGAAGAGTCTTATACTGACCGCTCTTCCCTGGCGCGAAGCGGATGAC 900
Db 841 AATTACAGCGTTGAAGAGTCTTATACTGACCGCTCTTCCCTGGCGCGAAGCGGATGAC 900
Qy 901 GTAAACAGAGCGGCGCAATGCCAAGCTCTTTTACCAATGGAGCGCTGATTCAAATTTGGCTG 960
Db 901 GTAAACAGAGCGGCGCAATGCCAAGCTCTTTTACCAATGGAGCGCTGATTCAAATTTGGCTG 960
Qy 961 TCGTCTTTGAAGCGGACTTCGATTATCAGAAACCAAGAGTGGCGCGGATT---AACAAA 1017
Db 961 TCGTCTTTGAAGCGGACTTCGATTATCAGAAACCAAGAGTGGCGCGGATT---AACAAA 1017
Qy 1021 GGCTCGTTCGCGAGGATTTCCACCTGGACGCGCACTATATCAGAAAGATTTCGAG 1080
Db 1018 GGTTCTGTTCCGAGGATTTACACCAATGGGAACCTGAGTACCAATAAAGGAAGTTGGC 1077
Qy 1081 AATATATACAAACCGCAGATGACACCGGATTTCAACGGTTTTTACTTTGCGGTATGACAGC 1140
Db 1078 GAAATATACAAACCGCAGATGACACCGGATTTCAACGGTTTTTACTTTGCGTTTTGGACAGC 1137
Qy 1141 CAACCGTTGCAACT---GGGCGGCGCAACATCGCTTGTGCTTAAACCTTTGCGCGAGTCGG 1197
Db 1138 CATCGTTGCAACTCGGCGGCGGCGGACACCGCTGCTGTTTAAACCTTTGCGCGAGCGC 1197
Qy 1198 CGTGAGTTGGAACCTTAAACCGGAGGATTTACTTTCAGCGAAGAGTATCCGCTACT 1257
Db 1198 CGTGAGTTGGAACCTTAAACCGGAGGATTTACTTTCAGCGGCGGCTGTTGTTGCAACC 1257
Qy 1258 ACCAGCTCGAATTCACACCCCGTGAACCACTAATATGTTTCTCACTGCTGTGATCAA 1317
Db 1258 ACCAGCAGTATCCAGCATCCGTTGAAACCAACCACTACGGTTTCTCACTGCTGACCAA 1317
Qy 1318 ATCAATGGAACGACGTTTCAGCAGCGGTGCGAGATATCCGTTACGATCATACCAAAATG 1377
Db 1318 ATCAATGGAACGACGTTTCAGTAGCGCGGAGGTATCCGTTACGACCAACCAAAATG 1377
Qy 1378 ACGCTCAGGAATTTGAATGCGAGTGTATGCTGTGCAAAACACCGCTGCGAGCCAAT 1437
Db 1378 ACGCTCAGGAATTTGAATGCGAGTGTATGCTGTGCAAAACACCACTGCGAGCCAAC 1437
Qy 1438 ACTTATAAGGCTGGAGCGGATTTGTCGGTTTGGCGGCGCAACTGAATCAGGCTTGGCAT 1497
Db 1438 ACTTATAAGGCTGGAGCGGATTTGTCGGTTTGGCGGCGCAACTGAATCAGGCTTGGCAT 1497
Qy 1498 GTCGGTTACGATTAATCTCCGGTACGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCT 1557
Db 1498 GTCGGTTACGATTAATCTCCGGTACGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCT 1557
Qy 1558 TACAACCAAGGCTTCGGGTAAATTTGCTGCCAATCCCAACTGGAAGCGGAGCGAGCAC 1617
Db 1558 TACAACCAAGGCTTCGGGTAAATTTGCTGCCAATCCCAACTGGAAGCGGAGCGAGCAC 1617
Qy 1618 ACCACACCTCTCTCAGAGCGCGCAGCAAAAGGTAATTTGGATGCGCAACCTGTAT 1677
Db 1618 ACCACACCTCTCTCAGAGCGCGCAGCAAAAGGTAATTTGGATGCGCAACCTGTAT 1677
Qy 1678 CAACCAATTTACCGCACTTCTCTGAGAGCAGAGCTGACCAACGAGGCGGATGTC 1737
```

```
Db 1678 CAACCAATTTACCGCAATTTCTCTGTAAGAGCAGAAGCTGACCAACGAGCGCACTCCC 1737
Qy 1738 GGCTGTACTCAGATGAATTAATACTACTAGGTATGTAAGCAATCTTATTCGAAAAACCG 1797
Db 1738 GGCTGTACTCAGGAAATGTTTACTACAGTATATGACGACCCCTTACAAAGAAAACTG 1797
Qy 1798 GAATGGCAGATGCAAAATATCGATAGGCCCGAATCCGTGCTCTTGAAGCTGACAGGCGGT 1857
Db 1798 GAATGGCAGATGCAAAATATCGATAGGCCCGAATCCGTGCTCTTGAAGCTGACAGGCGGT 1857
Qy 1858 CTGAATGTGCAAAAGTAGCGTCTTTTGTCTTGAAGGCTGGAATTTGTTGGCTGCTGCTG 1917
Db 1858 CTGAATGTGCAAAAGTAGCGTCTTTTGTCTTGAAGGCTGGAATTTGTTGGCTGCTGCTG 1917
Qy 1918 GGTATTCGAAAGACAACTGTCGCGGCAACAGAGCTGCTGTCACACAGCGCGCGAAA 1977
Db 1918 GGTATTCGAAAGACAACTGTCGCGGCAACAGAGCTGCTGTCACACAGCGCGCTGAAA 1977
Qy 1978 GTGATTTGCCGCTGTCGACTACGAAAGCCCGAGCGAAAAATGGGGTGTGTTCTCCCGCTG 2037
Db 1978 GTGATTTGCCGCTGTCGACTACGAAAGCTCCGAGCGAAAAATGGGGTGTGTTCTCCCGCTG 2037
Qy 2038 ACTTATCTGCGTGGAAAAAGGCGCAAGAGCGCGCAATACACCGTTTATGAAAAACAGGCG 2097
Db 2038 ACTTATCTGCGTGGAAAAAGGCTCAAGAGCGCGCAATACACCGTTTATGAAAAACAGGCG 2097
Qy 2098 CGGGGTACGCTTTTGCAGAAAAAGGTAAAGATTACCCGCTGGCTGAAACAAAGTCGGCTTAT 2157
Db 2098 TGGGTACGCTTTTGCAGAAAAAGGTAAAGATTACCCGCTGGCTGAAACAAAGTCGGCTTAT 2157
Qy 2158 GTGTTTGATATGACGGCTTCTACAACTGGCTTAAACCTGACTTTTGGCTGAGGCGTA 2217
Db 2158 GTGTTTGATATGACGGCTTCTACAACTGGCTTAAACCTGACTTTTGGCTGAGGCGTA 2217
Qy 2218 TATAATGTGTTCAACCGCAATACCACTTGGGATTCCTGCGCGGTTTGTATAGCTAC 2277
Db 2218 TACAACCTGTTTCAACCGCAATACCACTTGGGATTCCTGCGCGGTTTGTATAGCTAC 2277
Qy 2278 AGCACCAACCAACCGCTGACGAGATGCAAGGCTTAGACCGCTACCGCGCTCAGGC 2337
Db 2278 AGCACCAACCAACCGCTGACGAGATGCAAGGCTTAGACCGCTACCGCGCTCAGGC 2337
Qy 2338 CGTAATTTACCGCTGATCGCTGGAATGGAAGTTTGA 2373
Db 2338 CGTAATTTACCGCTGATCGCTGGAATGGAAGTTTGA 2373
```

## RESULT 8

AAT26998  
ID AAT26998 standard; cDNA; 2376 BP.

XX AC AAT26998;  
XX DT 27-AUG-2003 (revised)  
XX DT 25-MAR-2003 (revised)  
XX DT 28-AUG-1996 (first entry)  
XX XX  
XX DE N. meningitidis serotype A haemoglobin receptor cDNA.

XX XX Bacterial haemoglobin receptor; hmbR gene; antibody; vaccine; meningitis;  
XX XX ss.

XX OS Neisseria meningitidis serogroup A.

XX PN WO9612020-A2.

XX XX 25-APR-1996.

XX XX 17-OCT-1995; 95WO-US013623.

XX XX 18-OCT-1994; 94US-00326670.

XX PR 02-OCT-1995; 95US-00537361.

XX XX

Db 1681 TATCAAGCAATATACCGCAATTTCTGCTGAAGAGCAGAAGCTGACCACCGAGCGGAT 1740  
Qy 1735 GTCCGCTGTACTACAGATGATTAATCTACTACGATATGTGTAGCAATCTTATTCGCAAAA 1794  
Db 1741 GTACGCTGTACTACAGATGATTAATCTACTACGATATGTGTAGCAATCTTATTCGCAAAA 1800  
Qy 1795 CCGGAATGCGAGATGCAAAATATCGATAGGCGCCGAATCCGTGGTCTTGAGCTGACAGCG 1854  
Db 1801 CTGGAATGCGAGATGCAAAATATCGCAAGGCGCAATCCGGGTATCGAGCTGACGGCG 1860  
Qy 1855 COTCTGAATGTGACAAAAGTAGCGCTTTTGTCTTCTGAGGCTGGAATGTTTGGCTCG 1914  
Db 1861 CGTCTGAATGTGACAAAAGTAGCGCTTTTGTCTTCTGAGGCTGGAATGTTTGGCTCG 1920  
Qy 1915 CTGGGTATGCGAAAGCAAACTGTGCGGCGACACAGAGCTGCTGTCCACACAGCGCGG 1974  
Db 1921 CTGGGTATGCGAAAGCAAACTGTGCGGCGACACAGAGCTGCTGTCCACACAGCGGTG 1980  
Qy 1975 AAAGTGATTTGCGGTGTGACATCGAAGCCCGAGCGCAAAATGGGGTGTGTTCTCCCGC 2034  
Db 1981 AAAGTGATTTGCGGTGTGACATCGAAGCTGCAAGCTCGAGCGCAAAATGGGGTGTGTTCTCCCGC 2040  
Qy 2035 CTGACTTATCTGGGTGCGAAAGGCGCAAGAGCGCGCAATACACCGTTTATGAAACCAAG 2094  
Db 2041 CTGACCTATCTGGGCGCGAAAGGCTCAAGAGCGCGCAATACACCGTTTATGAAACCAAG 2100  
Qy 2095 GCGCGGGTACCGCTTTGCGAAGAAAGGTAAAGATTACCGTGGCTGAACAGTCGGCT 2154  
Db 2101 GCGTGGGTACCGCTTTGCGAAGAAAGGTAAAGATTACCGTGGCTGAACAGTCGGCT 2160  
Qy 2155 TATGTGTTTGTATGTACGGCTTCTACAACTGGCTTAAAGCTGACTTGGCTGCGAGCG 2214  
Db 2161 TATGTGTTTGTATGTACGGCTTCTACAACTGGCTTAAAGCTGACTTGGCTGCGAGCG 2220  
Qy 2215 GTATATATGTGTTTCAACGCGAAATACACCACTTGGGATTTCCCTGCGCGGTTGTATAGC 2274  
Db 2221 GTATATATGTGTTTCAACGCGAAATACACCACTTGGGATTTCCCTGCGCGGCTGTATAGC 2280  
Qy 2275 TACAGCACCAACCACTCGGTGCGACCGGATGCAAGGCTTAGACGCTTACCGCGCGCTCA 2334  
Db 2281 TACAGCACCAACCACTCGGTGCGACCGGATGCAAGGCTTAGACGCTTACCGCGCGCTCA 2340  
Qy 2335 GCGCGTAAATTACCGCTATCGCTGGAATGGAAAGTTTGA 2373  
Db 2341 ASCCGTAATTACCGCTATCGCTGGAATGGAAAGTTTAA 2379

## RESULT 7

AAT26997

AAT26997 standard; cDNA; 2373 BP.

XX AAT26997;

XX AC AAT26997;

XX DT 27-AUG-2003 (revised)

XX DT 25-MAR-2003 (revised)

XX DT 28-AUG-1996 (first entry)

XX XX

DE N. meningitidis serotype C haemoglobin receptor cDNA.

XX XX

KW Bacterial haemoglobin receptor; hmbR gene; antibody; vaccine; meningitis;

KW 88.

XX OS

XX Neisseria meningitidis serogroup C.

XX XX

PN WO9612020-A2.

XX XX

PD 25-APR-1996.

XX XX

XX 17-OCT-1995; 95WO-US013623.

XX XX

XX 18-OCT-1994; 94US-00326670.

PR 02-OCT-1995; 95US-00537361.

(UYOR-) UNIV OREGON HEALTH SCI.

XX Stojiljkovic I, So M, Hwa V, Heffron P, Nassif X;

XX WPI; 1996-222006/22.

DR P-PSDB; AAR95565.

XX DNA encoding Neisseria haemoglobin receptor proteins - for use in preparing polypeptide(s) and antibodies for vaccines for e.g. meningitis.

XX Claim 2; Page 46-49; 104pp; English.

XX A DNA sequence (AAT26997) codes for the bacterial haemoglobin receptor (HmbR) (AAR95565) of N. meningitidis serotype C isolate 8013 clone 6. It was isolated from a cosmid library by auxotroph complementation cloning and restriction enzyme digestion mapping of hemin utilisation positive cosmids. A 3.3 kb DNA fragment including the hmbR open reading frame was identified. The hmbR gene can be used to prepare recombinant receptor protein, useful as a vaccine for meningitis. It can be expressed in CC attenuated Salmonella host cells. It can also be utilised as a probe to CC detect infection in humans. Genes for N. meningitidis serotypes A and B CC and for N. gonorrhoeae were also obtd. (see also AAT26998-T27000).

XX (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 2373 BP; 636 A; 627 C; 610 G; 500 T; 0 U; 0 Other;

Query Match 83.2%; Score 1979.2; DB 2; Length 2373;

Best Local Similarity 90.2%; Pred. No. 0;

Matches 2142; Conservative 0; Mismatches 228; Indels 6; Gaps 2;

Qy 1 ATGAACCATTAACATGCTTCTTATTCGCGCGCTGGTTCGGCAGTATTTTCGGCAATCG 60

Db 1 ATGAACCATTAACATGCTTCTTATTCGCGCGCTGGTTCGGCAGTATTTTCGGCAATCG 60

Qy 61 GTCTTGGCAGCGATGAAGTGCACACCGGTTAAAGCAGAGATAAAGAA 120

Db 61 GTCTTGGCAGCGATGAAGTGCACACCGGTTAAAGCAGAGATAAAGAA 120

Qy 121 GTGCGGTTAAGACCACTTAATGCGGCTCAACCGTGGACGTTCAACCTCGGCGC 180

Db 121 GTGCGGTTAAGACCACTTAATGCGGCTCAACCGTGGACGTTCAACCTCGGCGC 180

Qy 181 ATCAACAGGAATGATACGGGCAACAAAGACTTGGTTCGCTTACTCCACCGACGTCGC 240

Db 181 ATCAACAGGAATGATACGGGCAACAAAGACTTGGTTCGCTTACTCCACCGACGTCGC 240

Qy 241 TTGAGCGATAGCGCGCCCATCAAAAGCTTTGTGTGCGCGCGGTGGAAGGCAACCGT 300

Db 241 TTGAGCGATAGCGCGCCCATCAAAAGCTTTGTGTGCGCGCGGTGGAAGGCAACCGT 300

Qy 301 GTGCGGTGACGATGACGGGTGAGCTGCTGATTCGGAGAAACTCAGCTGATGCA 360

Db 301 GTGCGGTGACGATGACGGGTGAGCTGCTGATTCGGAGAAACTCAGCTGATGCA 360

Qy 361 CGTTATGGCACTTCAACAGCTCGCGCTGTCTATCGACCCCGCACTCGTGGCAATAT 420

Db 361 CGTTATGGCACTTCAACAGCTCGCGCTGTCTATCGACCCCGCACTCGTGGCAATAT 420

Qy 421 GAAATCGAAGGCGCTGACCTTTTCAATACCGGTAGCGGCGATTTGGTGGCGCGGTG 480

Db 421 GAAATCGAAGGCGCTGACCTTTTCAATACCGGTAGCGGCGATTTGGTGGCGCGGTG 480

Qy 481 AATTACCAAACTTCGAGGACATGATTTGCTGTTGGACGACGACCAATTCGCGCGGTG 540

Db 481 AATTACCAAACTTCGAGGACATGATTTGCTGTTGGACGACGACCAATTCGCGCGGTG 540

Qy 541 ATGAAAAACGGTTACAGCAGCGCAACCGGCAATTCGCAAAATACACATCGTTTCGGTGTG 600

Db 541 ATGAAAAACGGTTACAGCAGCGGTAACCGTGAATGCAATACATCCCTCGTTTCGGTGTG 600

Qy 601 AGCAACGACCGCGTGGATGCGCGCTTTGTGTATTTCGCAACGTCGCGGTTCATGAGACCGAA 660

```

PR 18-OCT-1994; 94US-00326670.
XX 02-OCT-1995; 95US-00537361.
XX (UFOR-) UNIV OREGON HEALTH SCI.
XX Stojiljkovic I, So M, Hwa V, Heffron F, Nassif X;
XX WPI; 1996-222006/22.
XX P-PSDB; AAR95567.
XX
XX DNA encoding Neisseria haemoglobin receptor proteins - for use in
XX preparing polypeptide(s) and antibodies for vaccines for e.g. meningitis.
XX Claim 4; Page 58-61; 104pp; English.
XX
XX A DNA sequence (AAT26999) codes for the bacterial haemoglobin receptor
XX (hmbR) (AAR95567) of N. meningitidis serotype B. It was obtd. by PCR
XX amplification of genomic DNA using primers (AAT27003-04) based on the N.
XX meningitidis serotype C hmbR gene (see also AAT26997). The gene can be
XX used to prepare recombinant haemoglobin receptor protein, useful as a
XX vaccine for meningitis. It can be expressed in attenuated Salmonella host
XX cells. It can also be utilised as a probe to detect infection in humans.
XX hmbR genes for N. meningitidis serotype A (AAT26998) and for N.
XX gonorrhoeae (AAT27000) were similarly obtd. (Updated on 25-MAR-2003 to
XX correct PR field.) (Updated on 25-MAR-2003 to correct PI field.) (Updated
XX on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 2379 BP; 621 A; 632 C; 611 G; 515 T; 0 U; 0 Other;

Query Match 83.6%; Score 1988.6; DB 2; Length 2379;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 2149; Conservative 0; Mismatches 224; Indels 6; Gaps 2;

Qy 1 ATGAAACCATTTACACATGCTTCTATTGCGCGCTGCTGGCAGTATTTTCGGCAATCCG 60
Db 1 ATGAAACCATTTACAAATGCTCCCTATCCCGCGCTGCTGGCAGTATTTTCGGCAATCCG 60

Qy 61 GTCTTGCAGCGGATGAGCTGCAACCGAACACACCCCGTTAAAGCAGAGATAAAGAA 120
Db 61 GTCTTTCGGCAGATGAGCTGCAACTGAACTGAAACACACCCCGTTAAAGCAGAGTTAAAGCA 120

Qy 121 GTGCGCGTTAAAGACACAGCTTAATGCGCTGCAACCCGTTGAAACGCTGCAACCTCGCGCG 180
Db 121 GTGCGCGTTAAAGCGCAGCGCAATGCGCTGCGCTGTTGAAACGCTGCAACCTTAAACCGT 180

Qy 181 ATTCAACAGAAATGATACGCGCAACAAGACTTTGGTGGCTATTTCACCGATGTCGCG 240
Db 181 ATCAAAACAAGAAATGATACGCGCAACAAGACTTTGGTGGCTATTTCACCGATGTCGCG 240

Qy 241 TTGAGCGATAGCGCGCGCATCAAAAAGCTTTGCTGTGCGCGCGTGAAGGCAACCGT 300
Db 241 TTGAGCGATAGCGCGCGCATCAAAAAGCTTTGCTGTGCGCGCGTGAAGGCAACCGT 300

Qy 301 GTGCGGTGTGACGATGACGCGGTGAGCTGCTGATTTCGGAAGAAAATCACTGATGCA 360
Db 301 GTGCGGTGTGACGATGACGCGGTGAGCTGCTGATTTCGGAAGAAAATCACTGATGCA 360

Qy 361 CGTTATGGCAACTTCAACAGCTCGCGCTGCTATCGACCCCGAATCTGTCGCGCAACATC 420
Db 361 CGTTATGGCAACTTCAACAGCTCGCGCTGCTATCGACCCCGAATCTGTCGCGCAACATC 420

Qy 421 GAAATCGGAAGGCGCTGACTTTTCAATACCGGTAGCGCGCATTTGGGTGGCGCGTG 480
Db 421 GAAATCGGAAGGCGCTGACTTTTCAATACCGGTAGCGCGCATTTGGGTGGCGCGTG 480

Qy 481 AATTACCAAAACCTGCAAGGACATGATTGCTGTGACGACAGGCAATTCGGCGTGA 540
Db 481 AATTACCAAAACCTGCAAGGACATGATTGCTGTGACGACAGGCAATTCGGCGTGA 540

Qy 541 ATGAAAAACGTTTACAGACGCGCAACCGCGAATGGAACAAATACACTCGGTTTCGGTGTG 600
Db 541 ATGAAAAACGTTTACAGACGCGTAACCGTGAATGGAACAAATACACTCGGTTTCGGGCTG 600

601 AGCAACGACCGCGTGGATGCCCTTGTCTATTTCGCAACGTCGCGTCAATGACACGAA 660
601 AGCAACGACCGCGTGGATGCCCTTGTCTATTTCGCAACGTCGCGTCAATGACACGAA 660
661 AGCGCGGCGAGCGTGGCTATCCGCTAGAGGGTGTGGCAGCGGAGCAATATCCGTTGGT 720
661 AGCGCGGCGAGCGTGGCTATCCGCTAGAGGGTGTGGCAGCGGAGCAATATCCGTTGGT 720
721 TCGTACGCGGTATCCCTGATCCGTCACACACAAATACCAACCTTCTTGGGTAAAGATT 780
721 TCGTACGCGGTATCCCTGATCCGTCACACACAAATACCAACCTTCTTGGGTAAAGATT 780
781 GCTTATCAATCAACGACAGCACCAGCATCGGCGCATCGCTCAACGCTCAGCAGGGGCAT 840
781 GCTTATCAATCAACGACAGCACCAGCATCGGCGCATCGCTCAACGCTCAGCAGGGGCAT 840
841 AATTACAGATTTGAAGAGTCTTATACCTGACCGCTTCTTCTGCGCGGAGCGATGAC 900
841 AATTACAGATTTGAAGAGTCTTATCAACCTGCTTCTTATTGGCGTGAAGCTGACGAT 900
901 GTAAACAGAGCGCGCAATGCGCAACCTCTTTTACGAATGGACGCTGATTCAATTTGGCTG 960
901 GTAAACAGAGCGCGCAATGCGCAACCTCTTTTACGAATGGACGCTGATTCAATTTGGCTG 960
961 TCGTCTTTGAAGCGGAGTTCGATTATCAGACCAACCAAGTGGCGCGGTTAAACACAAA 1020
961 TCGTCTTTGAAGCGGAGTTCGATTATCAGACCAACCAAGTGGCGCGGTTAAACACAAA 1020
1021 GGCTCGTTTCCGA---CGGATTATTCCTCGAGCGCGCAACTATATATCAGAGGATTG 1077
1021 GGCTCGTTTCCGATAGAGGATCTTCCACCTTGACACGTAACATCAATCAAAAGGACTG 1080
1078 GAGATATATACACCGCAGCATGGACCCCGATTCAAAAGTCTTACTTTGCGGTATCGAC 1137
1081 GATGAATATCACAACCGCAGTATGGATACCCGCTTCAACGCAATACCTGCGGTTTGGAC 1140
1138 AGCCAAACCGTTGCAACT---GGCGCGCAACATCGCTTGTGCTTGTAAAACTTTTCGCCAGT 1194
1141 AGCCATCCGTTGCACTCGCGGCGGCGGACACCGCTGCTGTTTAAAACTTTTCGCCAGC 1200
1195 CGCGGTGAGTTTGAACCTTTAAACCGCGCAGATTTATCTTTCAGCGGAAAGATATCCGCT 1254
1201 CGCGGTGATTTGAAAACCTTAAACCGCGCAGATTTATCTTTCAGCGCGCGCTGTTGTCGA 1260
1255 ACTACAGCTCGATTCAACACCCGCTGAAACCACTAAATATGTTTCTCACTGCTCAT 1314
1261 ACCACAGCAGTATCCAGCATCCGCTGAAACCAACCACTACGCTTCTCACTGCTCAT 1320
1315 CAAATCCAATGGAAACGACGCTTTCAGCAGCGCTGAGATATCCGTTACGATCATACAAA 1374
1321 CAAATCCAATGGAAACGACGCTTTCAGTACGCGCGAGTATCCGTTACGATCATACAAA 1380
1375 ATGACGCTCAGGAATTTGAATGCGAGTGTCTGTTGTGACAAAACACCGCTTCGACGC 1434
1381 ATGACGCTCAGGAATTTGAATGCGAGTGTCTGTTGTGACAAAACACCGCTTCGACGC 1440
1435 AATACTTATAAGCTGAGCGGATTTGCTGCTTGGCGCGCAACTGAATCAGGCTTGG 1494
1441 AACATTTATAAGCTGAGCGGTTTGTGCTGCTTGGCGCGCACTGAATCAGGCTTGG 1500
1495 CATGTCGTTTACGACATTTACTTCCGCTACCGTGTCCCCAATGCGTCCGAAGTATTTTC 1554
1501 CGTGTGCTTACGACATTTACTTCCGCTACCGTGTCCCCAATGCGTCCGAAGTATTTTC 1560
1555 ACTTAAACCAACCGTTCCGGTAAATGCTGCCCAATCCCACTGAAGCGCGGCGCAGC 1614
1561 ACTTAAACCAACCGTTCCGGTAAATGCTGCCCAATCCCACTGAAGCGCGGCGCAGC 1620
1615 ACCACCCACACCTGCTCTGCAAGGCGCGCAGGAAAGAGTACTTTTGGATGCGCAACTG 1674
1621 ACCACCCACACCTGCTCTGCAAGGCGCGCAGGAAAGAGTACTTTTGGATGCGCAACTG 1680
1675 TATCAAAACAATTAACGCAACTTTTGTGTTGAAGAGCAGAGCTGACCAACGAGCGGCGAT 1734

```

```
Db 242056 ATGAAAAACGGTTACAGACCGGTAACCGTGAATGGACAAATACTCTCGGTTTCGGTGTG 241997
QY 601 AGCAACGCGCGTGGATCGCGTTTCTGTGTTATTCGACAGTTCGCGGTTCATCAGACCGNA 660
Db 241996 AGTAACGACCGCGTGGATCGCTCTTGTGTTATTCGACAGTTCGCGGTTCATGAAACCGNA 241937
QY 661 AGCGGGGCGAGCGTGGCTATCCGTTAGAGGTTGCTGGCAGCGAGCAATATTCGTTGT 720
Db 241936 AGTGGCGGNAACCGAGGCTATCTGTGGAAGGGAAGGAGTGGCGGNAATATCCGTTGT 241877
QY 721 TCGTCACGCGGTATCCCTGATCCGTCGTCACAAACAAATACCAAACTTCTTGGGTGAAT 780
Db 241876 TCGGACGCGGTATCCCTGATTCGTCACAAACAAATACCAAGCTTTTGGGTGAAT 241817
QY 781 GCTTATCAATCAACGACAGCACCGCATCGGCCATCGTTTAAAGCCGAGCGGGCAT 840
Db 241816 GCTTACCAATTAACGATTAACCAACCGCATCGGCCATCGTTTAAAGCCGAGCGGACAT 241757
QY 841 AATTACACGATTAAGAGTCTTATAACCTGACCGCTTCTTCTGGCGGAAAGCCGATGAC 900
Db 241756 AATTACACGTTGAAGAGTCTTATAACCTGACCGCTTCTTCTGGCGGAAAGCCGATGAC 241697
QY 901 GTAACAGACGGGCAATGCCAATCTTTTACGAATGACCGCTGATTCAAATTTGGGTG 960
Db 241696 GTAACAGACGGGCAATGCCAATCTTTTACGAATGATCGCTGATTCAAATTTGGGTG 241637
QY 961 TCGTCTTTGAAGCGGACTTCGATATATCAGACACCAAGTGGCGGTTAACACAA 1020
Db 241636 TCGTCTTTGAAGCGGACTTCGATATATCAGAAACCAAGTGGCGGTTAACACAA 241577
QY 1021 GCGTCTTCCGACGATTAATCCACTGGACCGCAATATAATCAGAAGATTTGGAG 1080
Db 241576 GCGTCTTCCGATGATTAATCCACTGGACCGCAATATAATCAGAAGATTTGGAG 241517
QY 1081 AATATATACACCGGAGATGACACCGGATTCAAAGCTTTTACCTTGGTATGACACAG 1140
Db 241516 GAAATATACACCGGAGATGACACCGGATTCAAAGCTTTTACCTTGGTATGACACAG 241457
QY 1141 CAACGTTGCAACT--GGGCGGCAACATCGCTTGTGCTTAAACTTTCCCGAGTCGG 1197
Db 241456 CATCGTTGCAACTCGGGGGGGGGGACACCGCTTGTGCTTAAACTTTCCCGAGTCGG 241397
QY 1198 CGTGAGTTTGAACCTTAAACCGGACGATTAATTTACCTCAGGGAAGATATCCCGTACT 1257
Db 241396 CGTGATTTTGAACCTTAAACCGGACGATTAATTTACCTCAGGCGCGTGTGTTGCGAAC 241337
QY 1258 ACCAGTCTGATTCACACCGGCTGAAACCACTAATATGTTTCTCACTGCTGATCA 1317
Db 241336 ACCAGGAGTATCCAGCATCCGCTGAAACCACTACGTTTCTCACTGCTGACCAA 241277
QY 1318 ATCCAATGAAACGAGTGTTCAGCAGCGCTGAGATATCGTTACGATCATACCAAAATG 1377
Db 241276 ATTCAATGAAACGAGTGTTCAGTACGCGGAGGATATCCGTTACGACCAACCAATG 241217
QY 1378 AGCCTCAGGAATTAATCCGAGTGTCTATGTTGACAAACACCGCTTCGACCCAAT 1437
Db 241216 AGCCTCAGGAATTAATCCGAGTGTCTATGTTGACAAACACCACTTCGACCCAAC 241157
QY 1438 ACTTATAAGGCTGAGCGGATTTGCTGGTTGGCGGCGCAACTGAATCAGGCTTGGCAT 1497
Db 241156 ACTTATAAGGCTGAGCGGATTTGCTGGCTTTGGCGGCGCAACTGAATCAGGCTTGGCAT 241097
QY 1498 GTCGGTTACGATTAATTTCCGCTACCGTTCCTCAATGCTGCGAGTGTATTTCACT 1557
Db 241096 GTCGGTTACGATTAATTTCCGCTACCGTTCCTCAATGCTGCGAGTGTATTTCACT 241037
QY 1558 TACAACACGTTTGGGTAATTTGGTTCGCCAATCCCAACTGAAAGCCGAGCGACACC 1617
Db 241036 TACAACACGTTTGGGTAATTTGGTTCGCCAATCCCAACTGAAAGCCGAGCGACACC 240977
QY 1618 ACCACACCTGTCTTCGACGCGCGGAGGAAAGGTAATTTGGATGCCAACCTGTAT 1677
Db 240976 ACCACACCTGTCTTCGACGCGCGGAGGAAAGGTAATTTGGATGCCAACCTGTAT 240917
```

## RESULT 6

AAT26999

ID AAT26999 standard; cDNA; 2379 BP.

XX

AC AAT26999;

XX

DT 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 28-AUG-1996 (first entry)

XX

DE N. meningitidis serotype B haemoglobin receptor cDNA.

XX

KW Bacterial haemoglobin receptor; hmbR gene; antibody; vaccine; meningitis; 58.

XX

QS Neisseria meningitidis serogroup B.

XX

PN WO9612020-A2.

XX

PD 25-APR-1996.

XX

PF 17-OCT-1995; 95WO-US013623.

XX

[illegible]

|   |  |
|---|--|
| RESULT 5  |  |
| AAF21611/C  |  |
| ID AAF21611 standard; DNA; 349980 BP.                                     |  |
| XX  |  |
| AC AAF21611;  |  |
| XX  |  |
| DT 13-MAR-2001 (first entry)  |  |
| XX  |  |
| DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:112.            |  |
| XX  |  |
| KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;   |  |
| KW diagnosis; antigen; detection; infection; gene therapy; antibacterial; |  |
| ds.   |  |
| XX  |  |
| OS Neisseria meningitidis.  |  |
| XX  |  |
| WO200066791-A1.   |  |
| XX  |  |
| PD 09-NOV-2000.   |  |
| XX  |  |
| PF 08-MAR-2000; 2000WO-US005928.  |  |
| XX  |  |
| PR 30-APR-1999; 99US-0132068P.  |  |
| PR 08-OCT-1999; 99WO-US023573.  |  |
| PR 28-FEB-2000; 2000GB-00004695.  |  |
| XX  |  |
| (CHIR ) CHIRON CORP.  |  |
| PA (GENO-) INST GENOMIC RES.  |  |
| XX  |  |
| PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC;                  |  |
| PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M,                 |  |
| PI Rappuoli R, Frazer CM, Grandi G;                                       |  |
| XX  |  |

DR  
 XX  
 PT  
 PT  
 XX  
 XX  
 PS  
 XX  
 XX  
 The present invention describes the full length genome of *Neisseria meningitidis* B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613 represent fragments of the NMB genomic sequence, as the sequence was too long to go in a record on its own it was split into 8 sequences which overlap each other at the beginning and end of each sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the *Neisseria* proteins given in AAB58550 to AAB59593, and AAF21589 to AAF21606 represent PCR primers which are used in the exemplification of the present invention. The NMB genome and fragments from it have antibacterial activity, and can be used in vaccines and gene therapy. *Neisseria* nucleic acids, proteins and/or antibodies which binds to the proteins can be used in compositions for treating or preventing infection due to *Neisseria* bacteria or as a diagnostic reagent for detecting the presence of *Neisseria* bacteria or antibodies raised to *Neisseria* bacteria. Computers, computer memory, computer storage medium or computer databases can be used in a search to identify open reading frames (ORFs) or coding sequences within the NMB genome. The DNA sequences provide further opportunities to find antigenic or immunogenic proteins which are more effective in vaccines than the outer membrane proteins currently used

Sequence 349980 BP; 87189 A; 93501 C; 84627 G; 84663 T; 0 U; 0 Other;  
 Query Match 85.7%; Score 2038.4; DB 3; Length 349980;  
 Best Local Similarity 91.4%; Pred. No. 0;  
 Matches 2172; Conservative 0; Mismatches 201; Indels 3; Gaps 1;

|    |        |  |        |
|----|--------|--|--------|
| Qy | 1      | ATGAAACCATTTACATGCTTTCATTTCGCGCGCTGTCGCGAGTATTTTCGGCAATCGG     | 60     |
| Db | 242596 | ATGAAACCATTTACAAATGCTCCCTATCGCCGCGTGTGTCGCGAGTATTTTCGGCAATCGG  | 242537 |
| Qy | 61     | GTCTTGGCAGCGGATGAAGCTGCACCGAAACACACCCGTTTAAAGCAGAGATAAAGAA     | 120    |
| Db | 242536 | GTCTTGGCAGCAGATGAAGCTGCAACTGAAACCAACCCGTTTAAAGCAGAGATAAAGCA    | 242477 |
| Qy | 121    | GTGCGCGTTTAAAGCCAGCTTAAATGCGCCTGCACCCGTGGAAACGTTCAACTCGGCGCG   | 180    |
| Db | 242476 | GTGCGCGTTTAAAGCTCAGCGCAATCGCGCTGCGGTGTGGAAACGCTCAACTTAACCGT    | 242417 |
| Qy | 181    | ATTCAACAGGAAATGATACGCGCAACAAAGACTTGGTGGCTTACTCCACCGACGTCGCGC   | 240    |
| Db | 242416 | ATCAACCAAGAAATGATACGCGCAATAAGACTTGGTGGCTTATCCACCGATGTCGCG      | 242357 |
| Qy | 241    | TTGAGCGATAGCGGCGCGCATCAAAAAGGCTTGTGTGCGCGCGTGGAAAGGCAACCGT     | 300    |
| Db | 242356 | TTGAGCGACAGCGCGCGCATCAAAAAGGCTTGTGTGCGCGCGTGGAAAGGCAACCGT      | 242297 |
| Qy | 301    | GTGCGTGTACGATTTGACGGCGTGAGCCTGCTGATTCGGAAGAAATCTACTGTATGCA     | 360    |
| Db | 242296 | GTGCGCGTGACCATAGACGGTGTAAACCTGCCTGATTCGAAGAAATCTGCTGTACGCGC    | 242237 |
| Qy | 361    | CGTTATGCAACTTCAACAGCTCGCGCCTGCTATTCGACCCCGAACTCGTGGCAACATC     | 420    |
| Db | 242236 | CGTTATGCAACTTCAACAGCTCGCGCTTGTCTATCGACCCCGAACTCGTGGCAACATC     | 242177 |
| Qy | 421    | GAAATTCGGAAGGCGCTGATCTTTTCAATACCGGTAGCGGCGATTCGGTGGCGCGCTG     | 480    |
| Db | 242176 | GAAATTCGGAAGGCGCGAGACTCTTTTCAATACCGGACGTCGTGTCATTCGGCGCGCTG    | 242117 |
| Qy | 481    | AATTACCAAAACCCCTGCAAGGACATGATTTGCTTTGGACGACAGGCAATTCGGCGCTGATG | 540    |
| Db | 242116 | AATTACCAAAACCCCTGCAAGGCGGATTTGCTTTGGACGACAGGCAATTCGGCGCTGATG   | 242057 |
| Qy | 541    | ATGAAAAACCGTTTACGACGCGCAACCCGGAATGGACAAATATCACTCGTTTCGGTGTG    | 600    |

```
QY 2338 CGTAATTACCGCGTATCGTGGATTGGAAGTTTCA 2373
Db |||||||||||||||||||||||||||||||||||
Db 105027 CGCAATTACCGCGTATCGTGGAAATGGAAGTTTAA 104992

RESULT 4
AAA81489_3/c
Continuation (4 of 9) of AAA81489 from base 300001 (N. meningitidis partial DNA sequence)
WP Sequence split into 9 fragments LOCUS AAA81489 Accession Aaa81489
WP Fragment Name Begin End
WP AAA81489_0 1 110000
WP AAA81489_1 100001 210000
WP AAA81489_2 200001 310000
WP AAA81489_3 300001 410000
WP AAA81489_4 400001 510000
WP AAA81489_5 500001 610000
WP AAA81489_6 600001 710000
WP AAA81489_7 700001 810000
WP AAA81489_8 800001 837096

Query Match 85.7%; Score 2038.4; DB 3; Length 110000;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 2172; Conservative 0; Mismatches 201; Indels 3; Gaps 1;

QY 1 ATGAAACCAATTACACATGCTTCTATTCGCGGCTGGTTCGGCAGTATTTTCGGCAATCG 60
Db |||||||||||||||||||||||||||||||||||
Db 7367 ATGAAACCAATTACAAATGCTCTATTCGCGGCTGGTTCGGCAGTATTTTCGGCAATCG 7308

QY 61 GTCTTGGCAGCGGATGAAGTGCACACCGAAACCAACCGGTTAAAGCAGAGATAAAAGAA 120
Db |||||||||||||||||||||||||||||||||||
Db 7307 GTCTTGGCAGCAGATGAAGTGCACACCGGTTAAAGCAGAGATAAAAGCA 7248

QY 121 GTGCGGTTAAAGCCGCTTAATGCGCTGCAACCGTGGACGTCTCAACTCGGCGCG 180
Db |||||||||||||||||||||||||||||||||||
Db 7247 GTGCGGTTAAAGGTCAGCGCAATGCGCTGCGGTGGAACCGGTCAACCTTAACCGT 7188

QY 181 ATTCAACAGGAATGATACGCGACCAAAAGACTTTGGTGCCTTACTCCACCGACGTCGCG 240
Db |||||||||||||||||||||||||||||||||||
Db 7187 ATCAACAGAGAAATGATACGCGACCAAAAGACTTTGGTGGCTATTTCCACCGATGTCGCG 7128

QY 241 TTGAGCGATPAGCGCGCGCATCAAAAGGCTTTGTGTGCGCGGTGGAAGGCAACCGT 300
Db |||||||||||||||||||||||||||||||||||
Db 7127 TTGAGCGACAGCGCGCGCATCAAAAGGCTTTGTGTGCGCGGTGGAAGGCAACCGT 7068

QY 301 GTCGGTGTGAGCTTGAAGCGGTGAGCTGCTGATTCGGAGAAATCTCACTGATGCA 360
Db |||||||||||||||||||||||||||||||||||
Db 7067 GTCGGCGTGAGCATAGACGGTGTAAACCTGCTGATTTCTGAAGAAATCTGCTGTACGCC 7008

QY 361 CGTTATGGCAACTTCAACAGCTCGCGCTGTCTATCGACCCCGAACTCGTGCGCAACATC 420
Db |||||||||||||||||||||||||||||||||||
Db 7007 CGTTATGGCAACTTCAACAGCTCGCGCTGTCTATCGACCCCGAACTCGTGCGCAACATC 6948

QY 421 GAAATCGCGAAGGCGCTGACTTTTCAATACCGTAGCGCGCATTTGGTGGCGCGGTG 480
Db |||||||||||||||||||||||||||||||||||
Db 6947 GAAATCGTGAAGGCGCGACTCTTTCAATACCGCGAGTGTGCAATTTGGGCGCGGTG 6888

QY 481 AATTACCAACCTCGAAGGACATGATTTGCTGTTGGACGACGACCAATTCGGCGTGATG 540
Db |||||||||||||||||||||||||||||||||||
Db 6887 AATTACCAACCTCGAAGGCGGTGATTTGCTGTTGGACGACGACCAATTCGGCGTGATG 6828

QY 541 ATGAAAAACGGTTACAGACGCGCAACCGCAATGGACAAATACACTCGTTTCGGTGTG 600
Db |||||||||||||||||||||||||||||||||||
Db 6827 ATGAAAAACGGTTACAGACGCGGTAAACCGTGAATGGACAAATACTCTCGTTTCGGTGTG 6768

QY 601 AGCAACGACCGCGTGAATCGCGCTTTGCTGATTCGCAACGTCGCGGTGATGAGACCGAA 660
Db |||||||||||||||||||||||||||||||||||
Db 6767 AGTAACGACCGCGTGAATCGCGCTTTGCTGATTCGCAACGTCGCGGTGATGAGACCGAA 6708

QY 661 ACGCGGCGCGCGTGGCTATCGGTAGAGGTGCTGCGAGCGGAGCAATATTCGTTGT 720
Db |||||||||||||||||||||||||||||||||||
Db 6707 AGTGGCGGAAACCGAGGCTATGCTGTGGAAGGGAAGGCGAGTGGCGCGGAATATCCGTTGT 6648
```

```
QY 721 TCGTACGCGGTATCCCTGATCCGTCAAACAACAATACCAAACTTCTTGGTAAAGATT 780
Db |||||||||||||||||||||||||||||||||||
Db 6647 TCGGACGCGGTATCCCTGATTCGTCAAACAACAATACCAAACTTCTTGGTAAAGATT 6588

QY 781 GCTTATCAAATCAACGACGACCGCATCGGCCATCGTTTAAAGCGCAGCAGGGGCAAT 840
Db |||||||||||||||||||||||||||||||||||
Db 6587 GCTTACCAAATTAACGATTAACCGCATCGGCCATCGTTTAAAGCGCAGCAGGGCAAT 6528

QY 841 AATTACACGATTGAAGAGTCTTATAAAGTCTGACCGCTTCTCTGCGGGAAGCCGATGAC 900
Db |||||||||||||||||||||||||||||||||||
Db 6527 AATTACACGATTGAAGAGTCTTATAAAGTCTGACCGCTTCTCTGCGGGAAGCCGATGAC 6468

QY 901 GTAAACAGACGCGCAATGCGCAACTCTTTTACGAATGACGCGCTGATTCAAATTTGGCTG 960
Db |||||||||||||||||||||||||||||||||||
Db 6467 GTAAACAGACGCGCAATGCGCAACTCTTTTACGAATGATGCTGATTCAAATTTGGCTG 6408

QY 961 TCGTCTTTGAAGCGGAGCTTCGATATCAGACCAACCAAGTGGCGGCGGTTAACACAAA 1020
Db |||||||||||||||||||||||||||||||||||
Db 6407 TCGTCTTTGAAGCGGAGCTTCGATATCAGACCAACCAAGTGGCGGCGGTTAACACAAA 6348

QY 1021 GGCTCGTTCCCGACGGATTATTCACCTGACGCGCAACTATAATCAGAAGGATTTGGAG 1080
Db |||||||||||||||||||||||||||||||||||
Db 6347 GGCTCGTTCCCGAGGATTAATTCACCTGACGCGCAACTATAATCAGAAGGATTTGGAG 6288

QY 1081 AATATATACAAACCGCAGCATGGACACCGGATTTCAAACGTTTTTACTTTGCGTATGACAGC 1140
Db |||||||||||||||||||||||||||||||||||
Db 6287 GAAATATACAAACCGCAGCATGGACACCGGATTTCAAACGTTTTTACTTTGCGTATGACAGC 6228

QY 1141 CAACGTTGCAACT---GGGCGGCGCAACATCGCTTGTGCTTAAACCTTTGCCAGTGG 1197
Db |||||||||||||||||||||||||||||||||||
Db 6227 CATCGTTGCAACTCGGCGGCGGCGCACACCGCTGTCTGTTTAAACCTTTTTCGTCAGCGC 6168

QY 1198 CGTCAGTTTGAATACTTAAACCGCGACGATTAATTAATTCAGCGAAAGATATCCCGTACT 1257
Db |||||||||||||||||||||||||||||||||||
Db 6167 CGTCAGTTTGAATACTTAAACCGCGACGATTAATTAATTCAGCGCGCGTGTGTTGCAACC 6108

QY 1258 ACCAGCTGATTTCAACACCCCGTGAATAACCACTAATATGTTTCTCACTGTGATCAA 1317
Db |||||||||||||||||||||||||||||||||||
Db 6107 ACCAGCAGTATCCAGCATCCGTTGAATAACCACTAATACGCTGTCTCACTGTGATCAA 6048

QY 1318 ATCAATGGAACGAGGTGTTTACGACGCGTGCAGATATCCGTTAGCATATACCAAAATG 1377
Db |||||||||||||||||||||||||||||||||||
Db 6047 ATCAATGGAACGAGGTGTTTACGAGTGCAGTATCCGTTAGCATATACCAAAATG 5988

QY 1378 ACGCTCAGGAATTAATGCGAGTGTCTGTCGAGTGTGACAAAACACCCCTGCGACCAAT 1437
Db |||||||||||||||||||||||||||||||||||
Db 5987 ACGCTCAGGAATTAATGCGAGTGTCTGTCGAGTGTGACAAAACACCACTGCGACCAAT 5928

QY 1438 ACTTATAAGGCTGAGCGGATTTGTCGGTTTGGCGCGCAACTGAATCAGGCTTGGCAT 1497
Db |||||||||||||||||||||||||||||||||||
Db 5927 ACTTATAAGGCTGAGCGGATTTGTCGGTTTGGCGCGCAACTGAATCAGGCTTGGCAT 5868

QY 1498 GTCGGTTACGACATTAATTCGGCTACCGTGTCCCAATGCTCGGAAGTGTATTTCACT 1557
Db |||||||||||||||||||||||||||||||||||
Db 5867 GTCGGTTACGACATTAATTCGGCTACCGTGTCCCAATGCTCGGAAGTGTATTTCACT 5808

QY 1558 TACAACACCGTTCGGGTAAATTTGCTGCGCAATCCAACTGAAAGCCGAGCGCAGCACC 1617
Db |||||||||||||||||||||||||||||||||||
Db 5807 TACAACACCGTTCGGGTAAATTTGCTGCGCAATCCAACTGAAAGCCGAGCGCAGCACC 5748

QY 1618 ACCCACACCTGTCTCTGCAAGGCGCGACGAAAAGGATCTTTGGATGCCAACCTGAT 1677
Db |||||||||||||||||||||||||||||||||||
Db 5747 ACCCACACCTGTCTCTGCAAGGCGCGACGAAAAGGATCTGATGCCAACCTGAT 5688

QY 1678 CAAAACAAATTAACGCACTTCTGTCAGAGCAGAGCTGACCAACGAGCGGCGATGTC 1737
Db |||||||||||||||||||||||||||||||||||
Db 5687 CAAAACAAATTAACGCACTTCTGTCAGAGCAGAGCTGACCAACGAGCGGCGATGTC 5628

QY 1738 GGCTGTACTCAGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1797
Db |||||||||||||||||||||||||||||||||||
Db 5627 GGCTGTACTCAGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5568

QY 1798 GAATGGCAGATGCAAAATATCGATAAGGCCCGCAATCCCGTGTCTTTGAGCTGACAGCCGT 1857
Db |||||||||||||||||||||||||||||||||||
```







Db 1921 TATCGAAGAAAGCAACTGTTCGGGCGACACAGCGCTGCTCCACACAGCGCGCGAAGTG 1980  
 Qy 1981 ATTGCCGGTGTGACTAGCAAGAGCCGAGCGAATAATGGGGTGTGTTCTCCGCTGACT 2040  
 Db 1981 ATTGCCGGTGTGACTAGCAAGAGCCGAGCGAATAATGGGGTGTGTTCTCCGCTGACT 2040  
 Qy 2041 TATCTGGTGTGCGAAGAGCCGAGCGCAATACACACCGTTTATGAAAACAAGGCGCGG 2100  
 Db 2041 TATCTGGTGTGCGAAGAGCCGAGCGCAATACACACCGTTTATGAAAACAAGGCGCGG 2100  
 Qy 2101 GGTAACGCTTTGCGAAGAGGTAAAGATTACCCGTGGCTGCAACAAAGTCGGCTTATGTG 2160  
 Db 2101 GGTAACGCTTTGCGAAGAGGTAAAGATTACCCGTGGCTGCAACAAAGTCGGCTTATGTG 2160  
 Qy 2161 TTTGATATGATACGGCTTCTACAACTGGCTTAAACCTGACTTTCGTGAGGCGGTATAT 2220  
 Db 2161 TTTGATATGATACGGCTTCTACAACTGGCTTAAACCTGACTTTCGTGAGGCGGTATAT 2220  
 Qy 2221 AATGTGTTCAACCGCAATACACCACTTGGGATTCCCTGCGCGTTTGTATAGCTACAGC 2280  
 Db 2221 AATGTGTTCAACCGCAATACACCACTTGGGATTCCCTGCGCGTTTGTATAGCTACAGC 2280  
 Qy 2281 ACCACCAACCGGTCGACCGAGATGCGAAGGCTTAGACCGCTACCGCGCTCAGGCGGT 2340  
 Db 2281 ACCACCAACCGGTCGACCGAGATGCGAAGGCTTAGACCGCTACCGCGCTCAGGCGGT 2340  
 Qy 2341 AATTACCGCTATCGCTGGATTGGAAGTTTGAATTC 2378  
 Db 2341 AATTACCGCTATCGCTGGATTGGAAGTTTGAATTC 2378

## RESULT 2

AAA81460  
 ID AAA81460 standard; DNA; 20844 BP.  
 AC  
 XX AAA81460;  
 XX  
 DT 04-DEC-2000 (first entry)  
 XX  
 DE N. meningitidis partial DNA sequence gnm\_8 SEQ ID NO:8.  
 XX  
 KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;  
 KW antigen; vaccine; diagnosis; infection; antibacterial; identification;  
 KW Meningococcus B; MenB; ds.  
 XX  
 OS Neisseria meningitidis.  
 XX  
 PN WO200022430-A2.  
 XX  
 PD 20-APR-2000.  
 XX  
 PF 08-OCT-1999; 99WO-US023573.  
 XX  
 PR 09-OCT-1998; 98US-0103794P.  
 PR 30-APR-1999; 99US-0132068P.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 XX  
 PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;  
 PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;  
 PI Rappuoli R, Pizza M;  
 XX  
 DR WPI; 2000-318079/27.  
 XX  
 PT Isolated nucleotide sequences of Neisseria meningitidis which can be used  
 PT in the diagnosis and treatment of N. meningitidis infection and other  
 PT Neisserial infections, for example, N.gonorrhoea.  
 XX  
 PS Claim 7; Page 320-326; 1760pp; English.  
 XX  
 CC The present invention describes methods of obtaining immunogenic proteins  
 CC from Neisseria genomic sequences. AAA81453 to AAA82414 represent

CC specifically claimed Neisseria meningitidis genomic DNA sequences;  
 CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA  
 CC sequences and their corresponding proteins; AAA81254 to AAA81259 and  
 CC AAA81304 to AAA81321 represent PCR primers used in the isolation of  
 CC Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent  
 CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all  
 CC used in the exemplification of the present invention. The nucleic acid  
 CC sequences, protein sequences, and antibodies against them, can be used in  
 CC the manufacture of a composition. The composition can be used as a  
 CC medicament (or in the manufacture of a medicament) for treating,  
 CC preventing or diagnosing infection due to Neisserial bacteria. For  
 CC example, some of the identified proteins could be components of vaccines  
 CC against Meningococcus B; against all serotypes; and/or against all  
 CC pathogenic Neisseriae. Identification of sequences from the bacterium  
 CC will also facilitate production of biological probes, particularly  
 CC organism-specific probes. Attempts to make efficacious Meningococcus B  
 CC vaccines have failed mainly due to antigen tolerance. Multivalent  
 CC vaccines have also been tried but none have successfully overcome  
 CC antigenic variability. The provision of further, complete sequences may  
 CC provide an opportunity to identify secreted or surface exposed proteins  
 CC that may be presumed targets for the immune system and which are not  
 CC antigenically variable or at least more conserved than other more  
 CC variable regions  
 XX  
 SQ Sequence 20844 BP; 5244 A; 5005 C; 5645 G; 4949 T; 0 U; 1 Other;  
 Query Match 85.7%; Score 2038.4; DB 3; Length 20844;  
 Best Local Similarity 91.4%; Pred. No. 0;  
 Matches 2172; Conservative 0; Mismatches 201; Indels 3; Gaps 1;  
 Qy 1 ATGAAACCATTTACACATGCTTCTCTATTGCGCGCTGTCGCGAGTATTTTCGCAATCCG 60  
 Db 6854 ATGAACCATTTACAAATGCTCCTATCGCGCGCTGTCGCGAGTATTTTCGCAATCCG 6913  
 Qy 61 GTCTTGCGAGCGGATGAAGCTGCAACGAAACACACACCGCTTAAAGCAGAGATAAAGAA 120  
 Db 6914 GTCTTGCGAGCAGATGAAGCTGCAACTGAACACACACCGCTTAAAGCAGAGATAAAGCA 6973  
 Qy 121 GTGCGGTTAAAGACACGAGCTTAATGCGCTGCAACCGTGGAACGCTGCAACCTCGGCGC 180  
 Db 6974 GTGCGGTTAAAGGTGAGCGCAATGCGCTGCGGTGAGAACGCTCAACCTTAACCGT 7033  
 Qy 181 ATTCAACAGGAATGATACGCGCAACAAAGACTTGTCGCTTACTCCACCGAGCTCGGC 240  
 Db 7034 ATCAACAAAGAAATGATACGCGCAATAAAGACTTGTCGCGCTTATTCACCGAGTGTGCGC 7093  
 Qy 241 TTGAGCGATAGCGCGCCCATCAAAAGGCTTTCGTGTCGCGCGCTGGAAGCGCAACCGT 300  
 Db 7094 TTGAGCGACAGCGCGCCCATCAAAAGGCTTTCGTGTTTCGCGCGCTGGAAGCGCAACCGT 7153  
 Qy 301 GTCGGTGTACAGATTGACGCGCTGAGCTTCCTCTGATTTCGGAAGAAACTCACTGTATGCA 360  
 Db 7154 GTCGGTGTAGCATAGACGCTGTAAACCTGCTGATTCTGAAGAAACTCGTGTATGCC 7213  
 Qy 361 CGTTATGGCAACTTCAACAGCTCGCGCTGTCTATCGACCCCGAATCTGTGCGCAACATC 420  
 Db 7214 CGTTATGGCAACTTCAACAGCTCGCGCTGTCTATCGACCCCGAATCTGTGCGCAACATC 7273  
 Qy 421 GAAATGCGAAGCGCGCTGACTCTTTCAATACCGGTAGCGCGCATTTGGTGGCGCGTG 480  
 Db 7274 GAAATGCGAAGCGCGCGACTCTTTCAATACCGGTAGCTGTGATTTGGCGCGGTGTG 7333  
 Qy 481 AATTACCAACCTGCAAGGACATGATTTGCTGTGGACGACAGGCAATTTTCGCGCTGATG 540  
 Db 7334 AATTACCAACCTGCAAGGCGCTGATTTGCTGTGGACGACAGGCAATTTTCGCGCTGATG 7393  
 Qy 541 ATGAAAAACGGTTACAGCGCGCAACCGCAATGAGCAAAATACATCTCGGTTTCGCTGTG 600  
 Db 7394 ATGAAAAACGGTTACAGCGCGCAACCGCTGAATGAGCAAAATACATCTCTCGGTTTCGCTGTG 7453  
 Qy 601 AGCAACGACCGCTGGATGCGCTTTGCTGTTTTCGCAACGCTCGCGCTCATGAGACCGAA 660  
 Db 7454 AGTAACGACCGCTGGATGCTGCTTTGCTGTTTTCGCAACGCTCGCGCTCATGAGACCGAA 7513

(Hmbr) (AAR95567) of *N. gonorrhoeae* strain MS11A. It was obtd. by PCR amplification of genomic DNA using primers (AAT27001-02) based on the *N. meningitidis* serotype C hmrB gene (see also AAR26997). The gene can be used to prepare recombinant haemoglobin receptor protein, useful as a vaccine for gonorrhea. It can be expressed in attenuated *Salmonella* host cells. It can also be utilised as a probe to detect infection in humans. CC hmrB genes for *N. meningitidis* serotype A (AAR26998) and B (AAR26999) were similarly obtd. (Updated on 25-MAR-2003 to correct PR field.) CC (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to correct OS field.)

XX SQ Sequence 2378 BP; 629 A; 618 C; 606 G; 525 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 2378; DB 2; Length 2378;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACACATTACATGCTTCTATTTGCGCGCTGGTGGCAGTATTTTCGGCAATCCG 60  
 DB 1 ATGAACACATTACATGCTTCTATTTGCGCGCTGGTGGCAGTATTTTCGGCAATCCG 60  
 QY 61 GTCTTGGCAGCGATGAAGCTGCAACCGAACCACACCCGTTAAAGCAGAGATAAGAA 120  
 DB 61 GTCTTGGCAGCGATGAAGCTGCAACCGAACCACACCCGTTAAAGCAGAGATAAGAA 120  
 QY 121 GTGCGGTTAAAGACCACTTAATGCGCTGCAACCGTGGAAACGTTCAACCTCGGCCG 180  
 DB 121 GTGCGGTTAAAGACCACTTAATGCGCTGCAACCGTGGAAACGTTCAACCTCGGCCG 180  
 QY 181 ATTCAACAGGAATGATACGCGAACAAAGACTTGGTGGTACTTCCACCGACGTCGCG 240  
 DB 181 ATTCAACAGGAATGATACGCGAACAAAGACTTGGTGGTACTTCCACCGACGTCGCG 240  
 QY 241 TTGAGCGATAGCGCGCCCATCAAAAGGCTTGTGTCGCGGGGTGGAAGCAACCGT 300  
 DB 241 TTGAGCGATAGCGCGCCCATCAAAAGGCTTGTGTCGCGGGGTGGAAGCAACCGT 300  
 QY 301 GTCGGTGTGAGCAATTCACCGCTGAGCTTGCCTGATTTCGGAAGAAACTCACTGTATG 360  
 DB 301 GTCGGTGTGAGCAATTCACCGCTGAGCTTGCCTGATTTCGGAAGAAACTCACTGTATG 360  
 QY 361 CGTTATGGCAACTTCAACAGCTCGCGCTGTCTATCGACCCCGAACTCGTGGCGCAACAT 420  
 DB 361 CGTTATGGCAACTTCAACAGCTCGCGCTGTCTATCGACCCCGAACTCGTGGCGCAACAT 420  
 QY 421 GAAATCGGAAGGGCGCTGACTTTCAATACCGGTAGCGGCGCATTTGGTGGCGCGCTG 480  
 DB 421 GAAATCGGAAGGGCGCTGACTTTCAATACCGGTAGCGGCGCATTTGGTGGCGCGCTG 480  
 QY 481 AATTACCAACCTGCAAGGACATGATTTGCTTGGACGACAGGCAATTCGGCGGTGATG 540  
 DB 481 AATTACCAACCTGCAAGGACATGATTTGCTTGGACGACAGGCAATTCGGCGGTGATG 540  
 QY 541 ATGAAAAACGGTTACAGCAGCGCAACCGCAATGACAAATACACTCGGTTTCGGTGTG 600  
 DB 541 ATGAAAAACGGTTACAGCAGCGCAACCGCAATGACAAATACACTCGGTTTCGGTGTG 600  
 QY 601 AGCAACGCGCGTGTGATCGGCTTGTCTGATTTTCGCAACGTCGCGTCAATGACCGGAA 660  
 DB 601 AGCAACGCGCGTGTGATCGGCTTGTCTGATTTTCGCAACGTCGCGTCAATGACCGGAA 660  
 QY 661 AGCGCGGCGAGCGTGGCTATCCGTTAGAGGGTGTGGCAGCGGAGCAATTAATCCGTTGT 720  
 DB 661 AGCGCGGCGAGCGTGGCTATCCGTTAGAGGGTGTGGCAGCGGAGCAATTAATCCGTTGT 720  
 QY 721 TCGTCAAGCGGTATCCCTGATCCGTTCCAAACACAAATACCAAACTTTCTTGGGTAGATT 780  
 DB 721 TCGTCAAGCGGTATCCCTGATCCGTTCCAAACACAAATACCAAACTTTCTTGGGTAGATT 780  
 QY 781 GCTTATCAATCAACGACGACCGCATCGGCCCATCGTTTAAACCGCGCAGCGGGGCAT 840  
 DB 781 GCTTATCAATCAACGACGACCGCATCGGCCCATCGTTTAAACCGCGCAGCGGGGCAT 840

QY 841 AATTACAGATTGAAGAGTCTTATAAAGCTGACCGCTCTTCTCTCGCGGGAAGCCCGATGAC 900  
 DB 841 AATTACAGATTGAAGAGTCTTATAAAGCTGACCGCTCTTCTCTCGCGGGAAGCCCGATGAC 900  
 QY 901 GTAAACAGACGCGCAATGCGCAACCTCTTTTACGAATGAGCGCTGATTTCAAATTTGGCTG 960  
 DB 901 GTAAACAGACGCGCAATGCGCAACCTCTTTTACGAATGAGCGCTGATTTCAAATTTGGCTG 960  
 QY 961 TCGTCTTTGAGAGCGGAGCTTGAATATCAGAACCAAGTGGCGGCGGTAAACACAA 1020  
 DB 961 TCGTCTTTGAGAGCGGAGCTTGAATATCAGAACCAAGTGGCGGCGGTAAACACAA 1020  
 QY 1021 GGTCTGTTCCGACCGGATTTATTCACCTGGAGCGCAACTATAATCAGAAGATTGGAG 1080  
 DB 1021 GGTCTGTTCCGACCGGATTTATTCACCTGGAGCGCAACTATAATCAGAAGATTGGAG 1080  
 QY 1081 AATATATACACCGCAGCATGAGACACCGGATTTCAAAGTTTTACTTTGGGTATGACACG 1140  
 DB 1081 AATATATACACCGCAGCATGAGACACCGGATTTCAAAGTTTTACTTTGGGTATGACACG 1140  
 QY 1141 CAACCGTTTGAACCTTAAACCGCGCAACATCGCTTGTGCTTAAACCTTTGCCAGTGGCGT 1200  
 DB 1141 CAACCGTTTGAACCTTAAACCGCGCAACATCGCTTGTGCTTAAACCTTTGCCAGTGGCGT 1200  
 QY 1201 GAGTTTGAACCTTAAACCGCGCAACATTTACTTTTACGAGAAAGATATCCCGTACTACC 1260  
 DB 1201 GAGTTTGAACCTTAAACCGCGCAACATTTACTTTTACGAGAAAGATATCCCGTACTACC 1260  
 QY 1261 AGCTCGATTCAACCCCGTGGAAACCACTAATTAATGTTTCTCACTGCTCTGATCAAAATC 1320  
 DB 1261 AGCTCGATTCAACCCCGTGGAAACCACTAATTAATGTTTCTCACTGCTCTGATCAAAATC 1320  
 QY 1321 CAATGGAACGAGCTGTTTACGAGCGCGTGGAAACCACTAATTAATGTTTCTCACTGCTCTGATCAAAATC 1380  
 DB 1321 CAATGGAACGAGCTGTTTACGAGCGCGTGGAAACCACTAATTAATGTTTCTCACTGCTCTGATCAAAATC 1380  
 QY 1381 CCTCAGGAATTTGAATGCGGAGTGTATGTTGTAACAAACACCGCTGACGCCAATATCT 1440  
 DB 1381 CCTCAGGAATTTGAATGCGGAGTGTATGTTGTAACAAACACCGCTGACGCCAATATCT 1440  
 QY 1441 TATAAAGGCTGAGCGGATTTGTCGGTTTGGCGGCGCAACTGAATCAGGCTTTGGCATGTC 1500  
 DB 1441 TATAAAGGCTGAGCGGATTTGTCGGTTTGGCGGCGCAACTGAATCAGGCTTTGGCATGTC 1500  
 QY 1501 GGTACGACATTTACTTTCGGCTTACCGTGTCCCAATGCGTCCGAAAGTGTATTTCACTTAC 1560  
 DB 1501 GGTACGACATTTACTTTCGGCTTACCGTGTCCCAATGCGTCCGAAAGTGTATTTCACTTAC 1560  
 QY 1561 AACACCGGTTGCGGTAAATTTGGCTGCCCAATCCAACTGAAAGCGGAGCGGAGCACACC 1620  
 DB 1561 AACACCGGTTGCGGTAAATTTGGCTGCCCAATCCAACTGAAAGCGGAGCGGAGCACACC 1620  
 QY 1621 GACACCTGTCTCTCAAGGCGGAGCGGAAAGTACTTTTGGATGCCAACCTGATCA 1680  
 DB 1621 GACACCTGTCTCTCAAGGCGGAGCGGAAAGTACTTTTGGATGCCAACCTGATCA 1680  
 QY 1681 AACAAATACCGCAACTTCTTGTCTGAGAGCAGAGCTGACCAACCGCGCGGATGTCGCG 1740  
 DB 1681 AACAAATACCGCAACTTCTTGTCTGAGAGCAGAGCTGACCAACCGCGCGGATGTCGCG 1740  
 QY 1741 TGTACTCAGATGAATTAATCTACTACGATAGCGGATTTCCGTTGCTTGGCTGAGCTGACGCGCGTCTG 1800  
 DB 1741 TGTACTCAGATGAATTAATCTACTACGATAGCGGATTTCCGTTGCTTGGCTGAGCTGACGCGCGTCTG 1800  
 QY 1801 TGGCAGATGCAAAATATCGATAGCGCGGATTTCCGTTGCTTGGCTGAGCTGACGCGCGTCTG 1860  
 DB 1801 TGGCAGATGCAAAATATCGATAGCGCGGATTTCCGTTGCTTGGCTGAGCTGACGCGCGTCTG 1860  
 QY 1861 AATGTGACAAAAGTAGCGTCTTTTGTCTGAGGCTGGAATTTGTTCCGCTCGCTGGGT 1920  
 DB 1861 AATGTGACAAAAGTAGCGTCTTTTGTCTGAGGCTGGAATTTGTTCCGCTCGCTGGGT 1920  
 QY 1921 TATGCGAAAAGCAAACTGTGCGGCGCAACAGCTGCTGTCTCCACAGCGCGGAAAGTG 1980

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2005, 14:52:34 ; Search time 1270 Seconds  
(without alignments)  
11084.366 Million cell updates/sec

Title: US-09-665-358-7  
Perfect score: 2378  
Sequence: 1 ATGAACCAATACACATGCT.....GATTGGAAGTTTGAATTC 2378

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_16Dec04:\*

1: Geneseqn1980s:\*

2: Geneseqn1990s:\*

3: Geneseqn2000s:\*

4: Geneseqn2001as:\*

5: Geneseqn2001bs:\*

6: Geneseqn2002as:\*

7: Geneseqn2002bs:\*

8: Geneseqn2003as:\*

9: Geneseqn2003bs:\*

10: Geneseqn2003cs:\*

11: Geneseqn2003ds:\*

12: Geneseqn2004as:\*

13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1          | 2378   | 100.0       | 2378   | 2     | AAT27000    |
| 2          | 2038.4 | 85.7        | 20844  | 3     | AAA81460    |
| C 3        | 2038.4 | 85.7        | 110000 | 3     | AAA81489 2  |
| C 4        | 2038.4 | 85.7        | 110000 | 3     | AAA81489 3  |
| C 5        | 2038.4 | 85.7        | 34980  | 3     | AAF21611    |
| 6          | 1988.6 | 83.6        | 2379   | 2     | AAT26999    |
| 7          | 1979.2 | 83.2        | 2373   | 2     | AAT26997    |
| 8          | 1846.4 | 77.6        | 2376   | 2     | AAT26998    |
| 9          | 840.8  | 35.4        | 852    | 10    | ABZ37973    |
| 10         | 833.8  | 35.1        | 864    | 10    | ABZ37953    |
| 11         | 344.8  | 14.5        | 348    | 10    | ABZ37945    |
| 12         | 139.4  | 8.4         | 222    | 10    | ABZ37965    |
| 13         | 118    | 5.0         | 558    | 3     | AAA82041    |
| 14         | 103.4  | 4.3         | 2427   | 10    | ABZ38258    |
| 15         | 78.6   | 3.3         | 2880   | 4     | AAD05317    |
| 16         | 78.2   | 3.3         | 2751   | 4     | AAD05364    |
| 17         | 74     | 3.1         | 3300   | 6     | ABS67377    |
| C 18       | 72     | 3.0         | 14652  | 3     | AAA81482    |
| C 19       | 72     | 3.0         | 110000 | 3     | AAA81489 1  |
| C 20       | 69.2   | 2.9         | 918    | 2     | AAQ22836    |

|      |      |     |        |    |             |
|------|------|-----|--------|----|-------------|
| C 21 | 68.4 | 2.9 | 1401   | 11 | ABD15654    |
| 22   | 68.4 | 2.9 | 1767   | 11 | ABD15812    |
| 23   | 68.4 | 2.9 | 3738   | 11 | ABD15760    |
| C 24 | 67.6 | 2.8 | 558    | 2  | AAQ22835    |
| 25   | 66.6 | 2.8 | 2736   | 10 | ABZ41276    |
| 26   | 65.4 | 2.8 | 2809   | 2  | AAQ55618    |
| 27   | 65   | 2.7 | 3035   | 6  | ABS67378    |
| 28   | 65   | 2.7 | 3286   | 2  | AAx78929    |
| 29   | 65   | 2.7 | 3286   | 6  | AAD32061    |
| 30   | 64.2 | 2.7 | 2748   | 5  | AAD19529    |
| C 31 | 64.2 | 2.7 | 69936  | 3  | AAA81479    |
| 32   | 64.2 | 2.7 | 110000 | 3  | AAA81490 04 |
| 33   | 64.2 | 2.7 | 34980  | 3  | AAF21607    |
| 34   | 63.8 | 2.7 | 7642   | 2  | AAV82020    |
| 35   | 63.6 | 2.7 | 2370   | 11 | ABD11836    |
| 36   | 62.2 | 2.6 | 3012   | 12 | ADL03151    |
| 37   | 61.4 | 2.6 | 2000   | 8  | ADA71938    |
| 38   | 60.8 | 2.6 | 2826   | 2  | AAT38071    |
| C 39 | 60.8 | 2.6 | 110000 | 2  | AAT42063 07 |
| 40   | 60.6 | 2.5 | 49617  | 4  | AAF28541    |
| C 41 | 60.4 | 2.5 | 110000 | 2  | AAT42063 06 |
| 42   | 59   | 2.5 | 3144   | 13 | ADT05700    |
| C 43 | 59   | 2.5 | 106645 | 13 | ADT05645    |
| 44   | 58.8 | 2.5 | 3318   | 13 | ADT05375    |
| C 45 | 58.6 | 2.5 | 5562   | 13 | ADT05480    |

ALIGNMENTS

RESULT 1  
AAT27000  
ID AAT27000 standard; cdna; 2378 BP.

XX AAT27000;

XX 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 28-AUG-1996 (first entry)

XX N. gonorrhoeae haemoglobin receptor cDNA.

XX Bacterial haemoglobin receptor; hmbR gene; antibody; vaccine; gonorrhoea; ss.

OS Neisseria gonorrhoeae.

XX Key Location/Qualifiers

FT CDS 1..2373

FT /\*tag= a

XX WO9612020-A2.

XX 25-APR-1996.

XX 17-OCT-1995; 95WO-US013623.

XX 18-OCT-1994; 94US-00326670.

XX 02-OCT-1995; 95US-00537361.

XX (UYOR-) UNIV OREGON HEALTH SCI.

XX Stojiljkovic I, So M, Hwa V, Heffron F, Nassif X;

XX WPI; 1996-222006/22.

XX P-PSDB; AAR95568.

XX DNA encoding Neisseria haemoglobin receptor proteins - for use in preparing polypeptide(s) and antibodies for vaccines for e.g. meningitis.

XX Claim 5; Page 64-66; 104pp; English.

XX A DNA sequence (AAT27000) codes for the bacterial haemoglobin receptor

541 ATGAAAAACGGTTACAGCAGCCGCAACCGCAATGGAACAAATACACTCGGTTTCGGTGTG 600  
|||||  
Db  
623 ATGAAAAACGGTTACAGCAGCCGCAACCGCAATGGAACAAATACACTCGGTTTCGGTGTG 682  
|||||  
Qy  
601 AGCAACGACCGCGTGGATGCGCTTGTGTTATTCGCAACGTCGGCGGTGATGAGACCGAA 660  
|||||  
Db  
683 AGCAACGACCGCGTGGATGCGCTTGTGTTATTCGCAACGTCGGCGGTGATGAGACCGAA 742  
|||||  
Qy  
661 AGCGCGGCGCGTGGATGCGCTTGTGTTATTCGCAACGTCGGCGGTGATGAGACCGAA 720  
|||||  
Db  
743 AGCGCGGCGCGTGGATGCGCTTGTGTTATTCGCAACGTCGGCGGTGATGAGACCGAA 802  
|||||  
Qy  
721 TCGTCACGCGGTATCCCTGATCCGTCACCAACCAAAATACCAAACTCTTCTGGGTGAAGAT 780  
|||||  
Db  
803 TCTGCGCGGTATTCCTGATCCGTCGCAACCAAAATACCAAACTCTTCTGGGTGAAGAT 862  
|||||  
Qy  
781 GCTTATCAAAATCAACGACCAACGCAATCGGCAATCGTTTAAACGGCGAGCAGGGGAT 840  
|||||  
Db  
863 GCTTATCAAAATCAACGACCAACGCAATCGGCAATCGTTTAAACGGCGAGCAGGGGAT 922  
|||||  
Qy  
841 AATTAACAGATTGAAGAGTCTTATACCTGACCGCTTCTTCTGCGCGGAGCCGATGAC 900  
|||||  
Db  
923 AATTAACAGATTGAAGAGTCTTATACCTGACCGCTTCTTCTGCGCGGAGCCGATGAC 982  
|||||  
Qy  
901 GTAAACAGACGCGCAATGCGCAACCTCTTTTACGAATGGACCGCTGATTCAAATTTGGCTG 960  
|||||  
Db  
983 GTCAACAGACGCGGTAAACCAACCTCTTTTACGAATGGACCGCGCAATTCGACCGTTG 1042  
|||||  
Qy  
961 TCGTCTTTGAAGCGGACTTCGATTTATACGACCAACCAAGAGTGGCGGGTTAAACCAAA 1020  
|||||  
Db  
1043 TCTATGGTAAAGCGGATGTCGATTTATCAAAACCAAAAGTATCTCGCGTCACTACAAA 1102  
|||||  
Qy  
1021 GCTGCTTCCCGACCGATTATTCACCTGGACCGCACTATAATCAGAAGGATTGGAG 1080  
|||||  
Db  
1103 GGTTCGTTCCCGACCGATTATTCACCTGGACCGCACTATAATCAGAAGGATTGGAG 1162  
|||||  
Qy  
1081 AATATATCAACCGCAGCATGACACCGCGATTCAAAACGTTTATCTTGGTATGACACAGC 1140  
|||||  
Db  
1163 GAAATCTATAACCGCAGCATGATACAACTTCAAAACGTTTATCTTGGTATGACACAGC 1222  
|||||  
Qy  
1141 CAAACGTTGCAACT--GGGCGGCCAAACATCGCTTGTGCTTAAACCTTTCGCCAGTGG 1197  
|||||  
Db  
1223 CATCGGTTGCAACTCGGCGGGGGGACACCGCTGTCGTTCAAAACCTTTCGCCAGTGG 1282  
|||||  
Qy  
1198 CGTGAAGTTGAAACCTTAAACCGCAGCATTTACTTTCAGCGCAAGAGTATCCCGTACT 1257  
|||||  
Db  
1283 CGTGAATTTGAAACCTTAAACCGCAGCATTTACTTTCAGCGCGCGTGTGTTGCGAAC 1342  
|||||  
Qy  
1258 ACCAGCTCGATTCAACACCGCGTGAACCACTAATTTATGTTTCTCACTGCTGATCAA 1317  
|||||  
Db  
1343 ACCAAGATATCCAGCATCCGCGTGAACCACTAGGTTTCTCGCTGTCGACCAA 1402  
|||||  
Qy  
1318 ATCCAATGAAACGATGTTTACGAGCGGTGAGATATCCGTTAGCATCATACCAAAATG 1377  
|||||  
Db  
1403 ATCCAATGAAACGATGTTTACGAGCGGTGAGATATCCGTTAGCATCATACCAAAATG 1462  
|||||  
Qy  
1378 AGCCTCAGGAATTTGAATGCGGAGTGTATGCTTGACAAACCGCCTCGACCCAT 1437  
|||||  
Db  
1463 AGCCTCAGGAATTTGAATGCGGAGTGTATGCTTGACAAACCGCCTCGACCCAT 1522  
|||||  
Qy  
1438 ACTTAAAGGCTGAGCGGATTTGTCGTTTGGCGCGCACTGAATCAGGCTTGGCAT 1497  
|||||  
Db  
1523 ACTTAAAGGCTGAGCGGATTTGTCGCTTGGCGCGCAGCTGAGCCAAATGCGGT 1582  
|||||  
Qy  
1498 GTCGGTTACGATTTACTTTCGGGTCACGTTGCCCAATGCGTCCGAAGTGTATTTCACT 1557  
|||||  
Db  
1583 TTGGGTTACGATGTGACCTCAGGTTTTCGGGTCGCCGAATGCGTCTGAAGTGTATTTCACT 1642  
|||||  
Qy  
1558 TACAACCGGTTTCGGGTAATTTGGCTGCGCAATCCCAACCTCAAAACCGCGAGCGACACC 1617  
|||||  
Db  
1643 TACAACCGGTTTCGGGTAATTTGGAGGCTTAATCTTAATTTGAAGCGAAGCGACGAC 1702  
|||||  
Qy  
1618 ACCCACACCTGTCTCTGCAAGCGCGCAGCAAAAGGATTTGATGATGCCAACCTGTAT 1677  
|||||

1703 ACCCACACCTGTCTTTCAGGGCGCGGACAAAGGACACTGATGCCAACCTGTAT 1762  
|||||  
Qy  
1678 CAAACAAATTAACGCAACTTCTTGTCTGAAGAGCAAGCTGACCAACGAGCGGATGTC 1737  
|||||  
Db  
1763 CAAACAAATTAACGCAACTTCTTGTCTGAAGAGCAAGCTGACCTGTGAGCGGACACCC 1822  
|||||  
Qy  
1738 GCGTGTACTCAGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1797  
|||||  
Db  
1823 GCGTGTACTCAGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1882  
|||||  
Qy  
1798 GAATGGCAGATGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1857  
|||||  
Db  
1883 GATGGCAGATGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1942  
|||||  
Qy  
1858 CTGAATGTGCAAAAGTAGCGTCTTCTGTTCTGAGGCTGGAATTTCTTCCGCTCGCTG 1917  
|||||  
Db  
1943 CTGAATGTGCAAAAGTAGCGTCTTCTGTTCTGAGGCTGGAATTTCTTCCGCTCGCTG 2002  
|||||  
Qy  
1918 GGTATGCAAAAGCAAACTGTCCGGCGCAACAGCCTGTCTCCACACAGCGCGCGAAA 1977  
|||||  
Db  
2003 GGTATGCAAAAGCAAACTGTCCGGCGCAACAGCCTGTCTCCACACAGCGCGCGAAA 2062  
|||||  
Qy  
1978 GTGATTCGCGGTGTCGACTACGAAGCCGAGCGGAAAAATGGGGTGTGTTCTCCGCGCTG 2037  
|||||  
Db  
2063 GTGATTCGCGGTGTCGACTACGAAGCCGAGCGGAAAAATGGGGTGTGTTCTCCGCGCTG 2122  
|||||  
Qy  
2038 ACTTATCTGGTTCGCAAAAGCCGCAACAGCGCAATACCGTTTATGAAACAAAGGCG 2097  
|||||  
Db  
2123 ACTTATCTGGTTCGCAAAAGCCGCAACAGCGCAATACCGTTTATGAAACAAAGGCG 2182  
|||||  
Qy  
2098 CGGGTACGCTTTTCGCAAAAGGTAAGATTAACCGCTGCTGCAACAAAGTCGGCTTAT 2157  
|||||  
Db  
2183 TGGGTACGCTTTTCGCAAAAGGTAAGATTAACCGCTGCTGCAACAAAGTCGGCTTAT 2242  
|||||  
Qy  
2158 GTGTTTGAATGTACGCTTCTCAAACTGGCTAAAAACCTGACCTTGGGTGCGAGCGTA 2217  
|||||  
Db  
2243 GTGTTTGAATGTACGCTTCTCAAACTGGCTAAAAACCTGACCTTGGGTGCGAGCGGTG 2302  
|||||  
Qy  
2218 TATATGTGTTCAACCGCAATACCACTTGGGATTTCCCTGCGGCTTGTGATAGCTAC 2277  
|||||  
Db  
2303 TACAACCTGTTCAACCGCAATACCACTTGGGATTTCCCTGCGGCTTGTGATAGCTAC 2362  
|||||  
Qy  
2278 AGCACCAACCGCGCTGCGAGATGCGCAAGGCTTAGACCGCTACCGCGCTCAGCG 2337  
|||||  
Db  
2363 AGCACCAACCGCGCTGCGAGATGCGCAAGGCTTAGACCGCTACCGCGCTCAGCG 2422  
|||||  
Qy  
2338 CGTAATTAACCGCTATCGCTGGAATTTGGAAGTTT 2371  
|||||  
Db  
2423 CGCAATTACCGCTATCGCTGGAATTTGGAAGTTT 2456  
|||||

## RESULT 10

AF133695  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

AF133695  
Neisseria meningitidis HmoO (hmr) gene, complete cds.  
AF133695.1  
GI:6692577  
Neisseria meningitidis  
Neisseria meningitidis  
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
Neisseriaceae; Neisseria.  
1 (bases 1 to 1500)  
Zhu, W., Hunt, D.J., Richardson, A.R. and Stojiljkovic, I.  
Use of heme compounds as iron sources by pathogenic neisseriae  
requires the product of the hmoO gene  
J. Bacteriol. 182 (2), 439-447 (2000)  
20096688  
10629191  
2 (bases 1 to 1500)  
Hunt, D.J., Richardson, A.R. and Stojiljkovic, I.

| QY         | 1555  | ACTTACCAACACCGGTTCCGGTGAATTTGGCTGCCCAATCCCACTCTGAAAGCCGAGCGCAGC | 1514            |
|------------|---|---|-----------------|
| Db         | 1725  | ACTTACCAACACCGGTTCCGGTGAATTTGGCTGCCCAATCCCACTCTGAAAGCCGAGCGCAGC | 1784            |
| QY         | 1615  | ACCACCACACCGTCTCTGCAAGGCCGAGCGAAGAAAGGTACTTTGGATGCCACCTG        | 1674            |
| Db         | 1785  | ACACCCACACCGCTCTCTGCAAGGCCGAGCGAAGAAAGGTACTTTGGATGCCACCTG       | 1844            |
| QY         | 1675  | TATCAAAACAATTACCGCAACTTCTTCTCTGAAGAGCAGAAGCTGACCAACAGCGCGCAT    | 1734            |
| Db         | 1845  | TATCAAGCAATTACCGCAATTTCTGTCTGAAGAGCAGAAGCTGACCAACAGCGCGCAT      | 1904            |
| QY         | 1735  | GTCCGCTGTACTCAGATGAATTACTACTACGGTATGTGTAGCAATCTTATCCGAAAAA      | 1794            |
| Db         | 1905  | GTACGCTGTACTCAGATGAATTACTACTACGGTATGTGTAGCAATCTTATCCGAAAAA      | 1964            |
| QY         | 1795  | CCGGAATGGCAGATGCAAAATATCGATAAGGCCCGGAATCCGTGGTCTTTGAGCTGACAGGC  | 1854            |
| Db         | 1965  | CTGGAATGGCAGATGCAAAATATCGACAAGGCCCGAATCCGCGGTATCGAGCTGACAGGC    | 2024            |
| QY         | 1855  | CGTCTGAATGTGACAAAGTAGCGTCTTTTGTCTCTGAGGGCTGGAATTTGTTCCGGCTCG    | 1914            |
| Db         | 2025  | CGTCTGAATGTGACAAAGTAGCGTCTTTTGTCTCTGAGGGCTGGAATTTGTTCCGGCTCG    | 2084            |
| QY         | 1915  | CTGGGTTATCGAAAAAGCAAACTGTCCGGCGCACAAACGCTGTGTCTCACACAGCCGCGC    | 1974            |
| Db         | 2085  | CTGGGTTATCGAAAAAGCAAACTGTCCGGCGCACAAACGCTGTGTCTCACACAGCCGCTG    | 2144            |
| QY         | 1975  | AAAGTGATTCGCGGTGTGCACTACGAAGGCCGAGCGAAGAAATGGGGTGTGTTCTCCCGC    | 2034            |
| Db         | 2145  | AAAGTGATTCGCGGTGTGCACTATGAAAGTCCGAGCGCAAAATGGGGTGTGTTCTCCCGC    | 2204            |
| QY         | 2035  | CTGACTTTATCTGGGTGCGAAAAAGGCCAAAGAGCGCGCAATACACCGCTTTATGAAACCAAG | 2094            |
| Db         | 2205  | CTGACTTTATCTGGGTGCGAAAAAGGCCAAAGAGCGCGCAATACACCGCTTTATGAAACCAAG | 2264            |
| QY         | 2095  | GGCCGGGGTACGCGCTTTGCGAAAAAGGTAAAAAGATTACCCGTGGCTGAAACAAGTCGGCT  | 2154            |
| Db         | 2265  | GGCTGGGGTACGCGCTTTGCGAAAAAGGTAAAAAGATTACCCGTGGCTGAAACAAGTCGGCT  | 2324            |
| QY         | 2155  | TATGTGTTTGAATGTACGGCTTCTACAACTGGCTAAAAACCTGACTTTGCGTGCAGGC      | 2214            |
| Db         | 2325  | TATGTGTTTGAATGTACGGCTTCTACAAACCCGCTGAAAAACCTGACTTTGCGTGCAGGC    | 2384            |
| QY         | 2215  | GTATATAATGTGTTCAACCGCAATACACCACTTTGGGATTTCCCTGCGCGGTTTGTATAGC   | 2274            |
| Db         | 2385  | GTATATAATGTGTTCAACCGCAATACACCACTTTGGGATTTCCCTGCGCGCGCTGTATAGC   | 2444            |
| QY         | 2275  | TACAGCACCAACACCGCGTTCGACCGAGATGGCAAGGCTTTAGACCGCTACCGCGCGCTCA   | 2334            |
| Db         | 2445  | TACAGCACCAACACCTCGGTTCGACCGCGATGGCAAGGCTTTAGACCGCTACCGCGCGCTCA  | 2504            |
| QY         | 2335  | GGCCGTAATTAACGCGGTATCGCTGGATTTGAAAGTTTGA                        | 2373            |
| Db         | 2505  | AGCCGTAATTAACGCGGTATCGCTGGAATGGAAGTTTAA                         | 2543            |
| RESULT 9   |   |   |                 |
| NMU40860   |   |   |                 |
| LOCUS      | 2568 bp   | DNA   | linear          |
| DEFINITION | Neisseria meningitidis hemoglobin receptor (hmbR) gene, complete cds.                 |   | BCT 13-NOV-1996 |
| ACCESSION  | U40860  |   |                 |
| VERSION    | U40860.1  | GI:1666680  |                 |
| KEYWORDS   | Neisseria meningitidis  |   |                 |
| SOURCE     | Neisseria meningitidis  |   |                 |
| ORGANISM   | Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Neisseria. |   |                 |
| REFERENCE  | 1 (bases 1 to 2568)   |   |                 |
| AUTHORS    | Stojiljkovic, I., Larson, J., Hwa, V., Anic, S. and So, M.                            |   |                 |
| TITLE      | Submitted (19-NOV-1995) Igor Stojiljkovic, Microbiology &                             |   |                 |
| JOURNAL    | Direct Submission   |   |                 |



```
source 1..2652
/organism="Neisseria meningitidis"
/mol_type="genomic DNA"
/strain="DNM2"
/db_xref="taxon:487"
165..2543
/gene="hmbR"
165..2543
CDS
/gene="hmbR"
/note="tonB-dependent hemoglobin receptor; phase variable
expression"
/codon_start=1
/transl_table=11
/product="hemoglobin receptor"
/protein_id="AAF19056.1"
/db_xref="GI:6606087"
/translation="MKPLQMLPVAALVGSIFGNPVFADEAETETTPVKAEVKAVRVK
GGRNPAVERVNLNRIKQEMIRNDKLDVRYSTDVGLSDSRHQKGFVAVGVEGNRVG
VSDGNLPDSEENSLYARYGNFSRSLSDPELVNIDIVKGADSFNTGSGALGGV
NYQTQGRDLLPFRQFGVMKNGYSTNRNWTNLGFGVSNDRVDAAALLYSQRRGHE
TESAGKRGYPVEGAGSGANIRGARGIPDPQSHKYHSLFKIAYOINDNHRIGASLNG
OQHNVTVESVNLASVREADDVNRNRTNLFEWTPESDRLSMVKADVVDYOKTKV
SAVNYKSGPPIEDSLTENYNOKDLDEIYNRMDTRPKRITLRLDLSHPLQLGGRHR
LSFKTPASRRDENLRDNYCFSGRVRTISSIQHPVKITNTGFSLSQIQNDYVPS
RAGIRYDHTMTPQELNAECHACDKTPPAANTYKWSGFGVGLAQLNAQWRVGYDITS
GYRPNASEVITYNHGSDMLPNLKAERSTHTLSLQGRSEKGTLDANLYQSLNV
NFLSEQKLITSGDVSTOMNYGNCNPSEKLEWOMONIDKARIRGIELTQSLNV
DKVAGFPEGKLFGLGYSKSLGSDNSLLSTOPLKVIAGIDYESPSEKGVFRSLT
VLGAKKVDQAVTYENKNGWTPLOKKYKDPFLNKSAYVDMYGFYKPKNLTIRAG
VTVNFRKITTWDSLRGLYSITNSVDNRDGLDRYRAPSRYNVAVSLEWKP"
1323..1331
/misc_feature
/gene="hmbR"
/note="the slipped strand mechanism, involving this poly G
tract, controls phase variable expression; the 'on'
configuration contains 9 G residues, the 'off'
configuration contains 10 G residues"

ORIGIN
Query Match 83.5%; Score 1985.4; DB 1; Length 2652;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 2147; Conservative 0; Mismatches 226; Indels 6; Gaps 2;
QY 1 ATGAAACCAATTACATGTTCTTCTATGTCGGCGTGTGCGGAGTATTTTGGCAATCCG 60
DB 165 ATGAAACCAATTACAAATGTCCTGTCGCGCGTGTGCGGAGTATTTTGGCAATCCG 224
QY 61 GTCTTGGCAGCGATGACGCTCAACCCGAAACACACCCGTTTAAAGCAGAGATAAAGNA 120
DB 225 GTCTTGGCAGCGATGACGCTCAACCTGAAACCAACACCCGTTTAAAGCAGAGATAAAGCA 284
QY 121 GTGCGCGTTTAAAGACGAGCTTAATGCGCTGCAACCGTGAACGCTCAACCTCGGCGCG 180
DB 285 GTACGGGTTTAAAGCCAGCGCAATGCGCTGCGGCTGTGGAACGCTCAACCTTAACCGT 344
QY 181 ATTCAACAGGAATGATACGACCAACAAAGACTTGGTGCTTACTTCCACCGACGTCGCG 240
DB 345 ATCAAAACAGGAATGATACGACCAACAAAGACTTGGTGCGCTATTTCACCGACGTCGCG 404
QY 241 TTGAGCGATAGCGCGCCCATCAAAAGCGTTTGTGTCGCGCGGTGGAAGGCAACCGT 300
DB 405 TTGAGCGAGCGCGCGCCCATCAAAAGCGTTTGTGTTTCGCGCGGTGGAAGGCAACCGT 464
QY 301 GTGCGGTGTGAGCATTTGACGCGGTGAGCGCTGCTGATTTCGGAAGAAACTCACTGTATGCA 360
DB 465 GTGCGGTGTGAGCATGAGCGCGTAAACCTGCTGATTTCGGAAGAAACTCACTGTATGCGC 524
QY 361 CATTATGCAACTTCAACAGCTGCGCGCTGTCTATGACCCCGAACTCGTGGCGCAACATC 420
DB 525 CGTTATGCAACTTCAACAGCTCGCGCTGTCTATGACCCCGAACTCGTGGCGCAACATC 584
QY 421 GAAATCGCAAGGGCGCTCACTCTTTTCATACGATAGCGGCGCATTTGGTGGCGCGGTG 480
DB 585 GACATCGTAAAGGGCGGCACTCTTTTCATACGCGAGCGCGCTTTGGGCGCGGTG 644
```

774 AGTAACAGCCGCGTGATGCTGCTTTGCTGTAATTCGCAACGGCGCGCCATGAAACCGAA 833  
 661 AGCGCGGCGAGCTGCGTATCCGGTAGAGGGTCTGCGAGCGAGCAATATATCCGTTGGT 720  
 834 AGCGCGGCGAACCCGCGGTATCCGGTAGAGGGTCCGGGTAAAGAAACGAATATCCGCGGT 893  
 721 TCCTCAGCGGTATCCCTGATCCGTCCAAAACACAAATACCAACATCTCTTTGGGTAAAGATT 780  
 894 TCCGCGCGGCGATCCCGGATCCGTCCAAACACAAATACCAACATCTCTTTGGGTAAAGATT 953  
 781 GCTTATCAATCAACGACGACGACGATCCGATCCGATCCGATCCGATCCGATCCGATCCGAT 840  
 954 GCTTATCAATCAACGACGACGACGATCCGATCCGATCCGATCCGATCCGATCCGATCCGAT 1013  
 841 AATTACAGATTGAAGAGTCTTAACTGACGACGCTCTCTCTGCGCGAGCGAGCGATGAC 900  
 1014 AATTACAGATTGAAGAGTCTTAACTGACGACGCTCTCTCTGCGCGAGCGAGCGATGAC 1073  
 901 GTAAACAGAGCGCGCAATGCCAACCCTCTTTTACGAATGGAGCGCTGATTCAAATTTGGCTG 960  
 1074 GTAAACAGAGCGCGCAATGCCAACCCTCTTTTACGAATGGAGCGCTGATTCAAATTTGGTTG 1133  
 961 TCCTCTTTGAAGCGGAGCTTCGATATCAGACAAACAAAGTGGCGCGGTAAACAAACA 1020  
 1134 TCCTCTTTGAAGCGGAGCTTCGATATCAGACAAACAAAGTGGCGCGGTAAACAAACAAG 1193  
 1021 GGCTCTTTCCGCGGATTTATCCACCTGGACGCGCACTATATCAGAGGATTTGGAG 1080  
 1194 GGCTCTTTCCGCGGATTTATCACCACATGGGAACCTGAGTACCAATAAAGGAAGTTGGC 1253  
 1081 AATATATCAACCGCAGATGACACCCGATTCAAACGTTTTACTTTTCGCTATGGACAGC 1140  
 1254 GAATATATCAACCGCAGATGACACCCGATTCAAACGTTTTACTTTTCGCTATGGACAGC 1313  
 1141 CAACGTTTGCACCT---GGGCGGCGCAACATCGCTTGTCTTAAACATTTTCGCGAGTCGG 1197  
 1314 CATCGTTTGCACCTCGGCGGCGGCGACACCGCTGTCTTTTAAACATTTTCGCGAGCGC 1373  
 1198 CGTCAGTTTGAACCTTTAAACCGGAGATTTATTTACTTTCAGCGAAAGAGTATCCGTTACT 1257  
 1374 CGTCAGTTTGAACCTTTAAACCGGAGATTTATTTACTTTCAGCGCGCGTGTGTTTCGAACC 1433  
 1258 ACCAGCTCGATTCACACCCCGTGGAAACCACTAATATGTTTCTCACTGTCTGATCAA 1317  
 1434 ACCAGCAGTATCCAGCATCCGCTGAAACCAACCACTAGCGTTTCTCACTGTCTGACCAA 1493  
 1318 ATCAATATGAACGAGCTGTTTCAGCAGCGCGTGACAGATTCGTTACGATCATACCAAAATG 1377  
 1494 ATTCATATGAACGAGCTGTTTCAGTACGCGCGCAGGTATCCGTTACGACCAACCAAAAATG 1553  
 1378 ACGCTCAGGAATTTGAATCCGAGTGTATGCTTGTGACAAACACCCGCTGACAGCCAAT 1437  
 1554 ACGCTCAGGAATTTGAATCCGAGTGTATGCTTGTGACAAACACCACTTCGACGCAAC 1613  
 1438 ACTTATAAGGCTGGAGCGGATTTGTCGGTTTGGCGGCGCAACTGAATCAGGCTTGGCAT 1497  
 1614 ACTTATAAGGCTGGAGCGGATTTGTCGGTTTGGCGGCGCAACTGAATCAGGCTTGGCAT 1673  
 1498 GTCCGTTACGACATTTACTTCGGCTACCGTGTCCCAATGCGTCCGAAGTGTATTTCACT 1557  
 1674 GTCCGTTACGACATTTACTTCGGCTACCGTGTCCCAATGCGTCCGAAGTGTATTTCACT 1733  
 1558 TACAACACCGGTTCCGGTAAATTTGCTGCGCAATCCCAACCTGAAAGCGGAGCGAGCAC 1617  
 1734 TACAACACCGGTTCCGGTAAATTTGCTGCGCAATCCCAACCTGAAAGCGGAGCGAGCAC 1793  
 1618 ACCACACCGCTGCTTCGAAAGCGGCGAGCGGAAAGGATTTGATGATGCAACCTGTAT 1677  
 1794 ACCACACCGCTGCTTCGAAAGCGGCGAGCGGAAAGGATTTGATGATGCAACCTGTAT 1853  
 1678 CAACACATTTACCGCACTTCTCTGAGAGCGAGGATGACCAACCGGCGGATGTC 1737  
 1854 CAACAGCAATTTACCGCAATTTCTCTGTAAGAGCAGAGCTGACCAACCGGCGCACTCCC 1913

1738 GGCTGTACTCAGATGAATTAATACTACTCGGTATGTAGCAATCTTATTCGAAAAACCG 1797  
 1914 GGCTGTACTCAGATGAATTAATACTACTCAGTATATGAGCGACCCCTACAAAGAAAACTG 1973  
 1798 GAATGGCAGATGCAAAATATCGAATAGCCCGAATCCGTGGTCTTTGAGCTGACAGGCGGT 1857  
 1974 GATTTGGCAGATGAATTAATACTACTCAGCAAGCCAGAAATCCGCGTATCGAGCTGACAGGCGGT 2033  
 1858 CTGAATGTGACAAAGTAGCGTCTTTTGTCTGAGGGCTGGAATTTGTTGGGCTCGCTG 1917  
 2034 CTGAATGTGACAAAGTAGCGTCTTTTGTCTGAGGGCTGGAATTTGTTGGGCTCGCTG 2093  
 1918 GGTATATCGAAAAAGCAAACTGTTCGGGCGACAAACGCTGTCTGCCACACAGCCGCCGAAA 1977  
 2094 GGTATATCGAAAAAGCAAACTGTTCGGGCGACAAACGCTGTCTGCCACACAGCCGCCGAAA 2153  
 1978 GTGATTCGCCGTTCGATACGAAAGCCGAGCGAAAAATGGGGTGTGTTCTCCCGCTG 2037  
 2154 GTGATTCGCCGTTCGATACGAAAGCCGAGCGAAAAATGGGGTGTGTTCTCCCGCTG 2213  
 2038 ACTTATCTGGGTGCGAAAAAGCCAAAGCGCGCAATACACCGTTTATGAAAAACAAGGC 2097  
 2214 ACTTATCTGGGTGCGAAAAAGCCAAAGCGCGCAATACACCGTTTATGAAAAACAAGGC 2273  
 2098 CGGGGTACGCTTTTCAGAAAAAGGTAAAGATTTACCCGCTGGCTGGAACAAAGTCGGCTTAT 2157  
 2274 TGGGTACGCTTTTCAGAAAAAGGTAAAGATTTACCCGCTGGCTGGAACAAAGTCGGCTTAT 2333  
 2158 GTGTTGATATGTAAGCTTCTCAAACTGGCTTAAACCTGACTTTTCGCTGACAGGCGTA 2217  
 2334 GTGTTGATATGTAAGCTTCTCAAACTGGCTTAAACCTGACTTTTCGCTGACAGGCGTA 2393  
 2218 TATAATGTGTTCAACCGCAATACACCACTGGGATTTCCCTGCGCGGTTGTATAGCTAC 2277  
 2394 TACAACCTGTTTCAACCGCAATACACCACTGGGATTTCCCTGCGCGGTTTATAGCTAC 2453  
 2278 AGACACCAACCAACCGGTTCGACCGAGATGGAAGGCTTAGACCGCTACCGCGCTCAGGC 2337  
 2454 AGACACCAACCAACCGGTTCGACCGAGATGGAAGGCTTAGACCGCTACCGCGCTCAGGC 2513  
 2338 CGTAATTAACCGCTGATCGCTGGATTTGAAGTTTGA 2373  
 2514 CGCAATTAACCGCTGATCGCTGGATTTGAAGTTTGA 2549

## RESULT 8

AF105339  
 LOCUS  
 DEFINITION  
 Neisseria meningitidis hemoglobin receptor (hmbR) gene, complete cds.  
 ACCESSION  
 AF105339  
 VERSION  
 AF105339.1  
 KEYWORDS  
 Neisseria meningitidis  
 Neisseria meningitidis  
 Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
 Neisseriaceae; Neisseria.  
 ORGANISM  
 1 (bases 1 to 2652)  
 Lewis, L.A., Gibson, M., Hartman, K., Ownbey, T., Vaughn, J. and Dyer, D.W.  
 Phase variation of HpuAB and HmbR, two distinct haemoglobin receptors of *Neisseria meningitidis* DNM2  
 Mol. Microbiol. 32 (5), 977-989 (1999)  
 99291064  
 MEDLINE  
 10361300  
 2 (bases 1 to 2652)  
 Lewis, L.A., Gibson, M., Hartman, K., Vaughn, J. and Dyer, D.W.  
 Direct Submission  
 Submitted (10-NOV-1998) Microbiology and Immunology, University of Oklahoma Health Sciences Center, 940 S.L. Young Blvd., Oklahoma City, OK 73104, USA  
 JOURNAL  
 FEATURES  
 Location/Qualifiers

```

Db 2090 ACCACCCACACCTCTCTCTGCAAGGCGCAGCGAAAGGTTACTTTGGATGCCAACCTG 2149
Qy 1675 TATCAAAACAAATACCGCAACTTCTGTCTGAAGAGCAGAACTGACCAACGAGCGCGAT 1734
Db 2150 TATCAAGCAATATACCGCAATTTCTGTCTGAAGAGCAGAACTGACCAACGAGCGCGAT 2209
Qy 1735 GTCCGCTGTACTCAGATGAATTAATCTACTACGCTGTGTGTAGCAATCTTTATTCGGAATA 1794
Db 2210 GTACGCTGTACTCAGATGAATTAATCTACTACGCTGTGTGTAGCAATCTTTATTCGGAATA 2269
Qy 1795 CCGAATGCGAGATGCAAAATATCGAATAAGCCCGAATCCGCTGTGTAGCTGACAGCG 1854
Db 2270 CTGGAATGCGAGATGCAAAATATCGAATAAGCCCGAATCCGCTGTGTAGCTGACAGCG 2329
Qy 1855 CCTCTGAATGTCACAAAGTACGCTTTTCTTCTGAGGCTGGAATTTGCGCTCG 1914
Db 2330 CCTCTGAATGTGACAAAGTACGCTTTTCTTCTGAGGCTGGAATTTGCGCTCG 2389
Qy 1915 CTGGGTTATGCGAAAGCAAACTGTCCGCGCAACAGCCTGCTGTCCACACAGCGCGCG 1974
Db 2390 CTGGGTTATGCGAAAGCAAACTGTCCGCGCAACAGCCTGCTGTCCACACAGCGCGTTG 2449
Qy 1975 AAAGTGAATGCGGCTGTGCACTACGAAAGCCGAGCGAAATGCGGCTGTCTCCCGC 2034
Db 2450 AAAGTGAATGCGGCTGTGCACTATGAAAGTCCGAGCGAAATGCGGCTGTCTCCCGC 2509
Qy 2035 CTGACTTATCTGGTTCGAAAGCCAAAGCCAAAGAGCGGCAATACACCTTTATGAAACAG 2094
Db 2510 CTGACTTATCTGGTTCGAAAGCCAAAGAGCTCAAGAGCGGCAATACACCTTTATGAAACAG 2569
Qy 2095 GCGCGGGTACGCTTTGCGAAGAAAGTAAAGATTACCGCTGGCTGAAACAGTCCGCT 2154
Db 2570 GCGTGGGTACGCTTTGCGAAGAAAGTAAAGATTACCGCTGGCTGAAACAGTCCGCT 2629
Qy 2155 TATGTGTTTGAATGATGACGCTTCTACAACTGGCTTAAAGAACTGATTTGCGTCCAGC 2214
Db 2630 TATGTGTTTGAATGATGACGCTTCTACAACTGGCTTAAAGAACTGATTTGCGTCCAGC 2689
Qy 2215 GTATATATGCTTCAACCGCAATACACCACTTGGGATTCCTGCGGCTGTGTATAGC 2274
Db 2690 GTATATATGCTTCAACCGCAATACACCACTTGGGATTCCTGCGGCTGTGTATAGC 2749
Qy 2275 TACAGACCAACCAACGCGCTGACGAGATGCAAGAGCTTACGCGCTACCGCGCTCA 2334
Db 2750 TACAGACCAACCAACGCGCTGACGAGATGCAAGAGCTTACGCGCTACCGCGCTCA 2809
Qy 2335 GCGCTGAATACGCGCTATGCTGTGATGGAAGTTTGA 2373
Db 2810 AGCGCTAATACGCGCTATGCTGTGATGGAAGTTTGA 2848

```

## RESULT 7

```

NMU40859
LOCUS NMU40859 2632 bp DNA linear BCT 13-NOV-1996
DEFINITION Neisseria meningitidis hemoglobin receptor (hmbR) gene, complete cds.

```

```

ACCESSION U40859
VERSION U40859.1 GI:1666678

```

```

KEYWORDS
SOURCE

```

```

ORGANISM

```

```

Neisseria meningitidis

```

```

Neisseria meningitidis

```

```

Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;

```

```

Neisseriaceae; Neisseria.

```

```

1 (bases 1 to 2632)

```

```

Stojiljkovic, I., Larson, J., Hwa, V., Anic, S. and So, M.

```

```

Direct Submission

```

```

Submitted (19-NOV-1995) Igor Stojiljkovic, Microbiology &

```

```

Immunology, OHSU, L220, Portland, OR 97201, USA

```

```

Location/Qualifiers

```

```

1..2632

```

```

/organism="Neisseria meningitidis"

```

```

/mol type="genomic DNA"

```

```

/strain="Serogroup A"

```

```

FEATURES
source

```

```

gene
CDS

```

```

/db_xref="taxon:487"
174..2549
/feature="hmbR"
174..2549
/feature="hmbR"
/feature="hmbR"
/feature="hmbR"
/codon_start=1
/transl_table=11
/product="hemoglobin receptor"
/protein_id="AAB18728.1"
/db_xref="GI:1666678"
/translation="MKQLQMLPIALVGSIFGNPVLAADEAETTPVKAEIKAVRVK
GORNAPAEVRLNRIKIQMLRDKNLVRYSTDLGSDSRGHRGKGFVAVGEGNRVG
VSDGNLPLDSEENSLVARYGNENSRSLIDPELVNTEIKVGDASFNFGSALGGV
NYOTLQGRDILLDDORQFCVMKNGYSTENBKTWTLGEGVNDVRDAALLYSORGE
TESARGYFVEGAGKETIRGASRGIFDPSKHKHNFGLKIAQIYNHRIASLNG
OQGHNTVEESYNLTASRRBDDVNRNRNANLFYEMWPDNSLWSSLKADFDYQKTKV
AAINTGLFTNYTWTETHEHKEVEGTINRSMDFRFTLLRLDLSHLQGLGHRRL
SPKTFASRRDFENLRNDDYFSGRVVTTSSIQHPVKTNTYFSLDQIQWMDVTSR
AGIRYDHTKTPQELNAEACADKTPPAANTYKWSGFVGLAAQLQAWRVGYDITSG
YRVPASVETYNHSGSNWLPNPKAERSTHTLSLQGRSEKGLDANLYOSVRYN
FLSEBQKLTSGTFCSTENAYYSICSDPYKEKLDWMKNIDKARIETLGRNLVD
KVASFVPGWKLFGSLKSKSDNLSLSTQPLKVIAGIDIESPSKMGVFSRLTY
LGAKVKDAQYTVYENKMGWTFLOKKVXDYPLNKSAYFDMYGFYKVNLTLAGV
YNLFRKYTTWDSLRGLYSYSTTNAVDKGLDRYRAPGRNYAVSLEWKF"

```

## ORIGIN

```

Query Match 83.6%; Score 1987.2; DB 1; Length 2632;
Best Local Similarity 90.1%; Pred. No. 0;
Matches 2140; Conservative 0; Mismatches 233; Indels 3; Gaps 1;

Qy 1 ATGAACCAATTACACATGCTTCCTATTGCGCGCTGTCGCGAGTATTTTCGCGCAATCCG 60
Db 174 ATGAACCAATTACAAATGCTCCCTATCGCGCGCTGTCGCGAGTATTTTCGCGCAATCCG 233

Qy 61 GTCTTGGCAGCGGATGAAGCTGCAACCGAAACCAACACCGCTTAAAGCAGAGATAAAGAA 120
Db 234 GTCTTGGCAGCAGATGAAGCTGCAACTGAACCAACCAACCGCTTAAAGCAGAGATAAAGCA 293

Qy 121 GTGCGCGTTAAAGACAGCTTAATCGCGCTGCAACCGTGGAAACGTTGCAACCTCGCGCGC 180
Db 294 GTGCGCGTTAAAGGTCAGCGCAATCGCGCTGTCGCGCTGTGGAACGCGCTTAACCGT 353

Qy 181 ATTCAACAGGAAATCATACGCGACAAACAAAGACTTTGTCGCTTACTCCACCGACGTCGCG 240
Db 354 ATCAACAGAAATCATACGCGACAAATAGACTTTGTCGCGCTTACTCCACCGATGTCGCG 413

Qy 241 TTGACCGATAGCGCGCGCATCAAAAAGCTTTGTCGCGCGCTGGAAGCGCAACCGT 300
Db 414 TTGACCGACAGCGCGCGCATCAAAAAGCTTTGTCGCGCGCTGGAAGCGCAACCGT 473

Qy 301 GTCGCTGTGAGCATTTGACGCGGTGAGCTGCTGATTCGGAAGAAACTCACTGTATGCA 360
Db 474 GTCGCGGTGAGCATAGACGCTGTAACCTGCTGATTCGGAAGAAACTCGCTGTACGCG 533

Qy 361 GCTTATGCAACTTCAACAGCTCGCGCTGCTATGACGCCCGCAACTCGTCGCGCAACATC 420
Db 534 CGTTATGCAACTTCAACAGCTCGCGCTTGTCTATGACGCCCGCAACTCGTCGCAATAT 593

Qy 421 GAAATCGCAAGGCGCTGACTTTTCAATACCGGTAGCGCGCAATTTGGTTCGCGCGCTG 480
Db 594 GAAATCGTGAAGGCGCAGACTCTTTTCAATACCGCAGTGGTGATTTGGCGCGCGGTG 653

Qy 481 AATTACCAACCCCTCAAGGACATGATTTGCTGTTGGACGACGAGCAATTCGCGCGTGATG 540
Db 654 AATTACCAACCGCTCAAGGCGCTGATTTGCTGTTGGACGACGAGCAATTCGCGCGTGATG 713

Qy 541 ATGAAAAACGTTTACAGCAGCGCAACCGCAATGGAACAAATACACTCGGTTTCGCTG 600
Db 714 ATGAAAAACGTTTACAGCAGCGGTAAACCGTGAATGGAACAAATACCTCGGTTTCGCTG 773

Qy 601 AGCAACGACCGCGTGGATGCGCTTTTGTGTTATTCGCAACGTCGCGGTTCATGAGACCGAA 660

```

Db 2690 GTATATAATGTTTCAACCGCAATACACCACTGGGATTCCTTCGCGGCTGTATAGC 2749  
Qy 2275 TACAGCACCAACCGGTCGACCGAGATGGCAAGCTTAGACCGCTACCGGCGCTCA 2334  
Db 2750 TACAGCACCAACCACTCGGTCGACCGGATGGCAAGCTTAGACCGCTACCGGCGCTCA 2809  
Qy 2335 GGCGTAAATACGCGGTATCGCTGGATTTGA 2373  
Db 2810 AGCGTAATTACGCGGTATCGCTGGAATGAAGTTTAA 2848

RESULT 6  
LOCUS I85729 3318 bp DNA linear PAT 10-JUN-1998  
DEFINITION Sequence 1 from patent US 5698438.  
ACCESSION I85729  
VERSION I85729.1 GI:3205447  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 3318)  
AUTHORS Scojliljkovic, I., So, M., Hwa, V., Heffron, P. and Nassif, X.  
TITLE Bacterial hemoglobin receptor gene  
JOURNAL Patent: US 5698438-A 1 16-DEC-1997;  
FEATURES Location/Qualifiers  
source 1..3318  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 83.6%; Score 1988.6; DB 6; Length 3318;  
Best Local Similarity 90.3%; Pred No. 0;  
Matches 2149; Conservative 0; Mismatches 224; Indels 6; Gaps 2;  
Qy 1 ATGAACCAATTACACATGCTTCTTATTCGCGGCTGTCGCGAGTATTTTCGCAATCCG 60  
Db 470 ATGAACCAATTACAAATGCTCCCTATCGCGGCTGTCGCGAGTATTTTCGCAATCCG 529  
Qy 61 GTCTTGCGAGGATGAAGCTGCAACCGCAACACACCGCTTAAGCAGAGATAAAGAA 120  
Db 530 GTCTTGCGGAGATGAAGCTGCAACCGCAACACACCGCTTAAGCAGAGATAAAGAA 589  
Qy 121 GTGCGGCTTAAAGCAGCTTAAATGCGCTGCAACCGTGAACCTGTCAACCTCGGCGC 180  
Db 590 GTGCGGCTTAAAGCAGCTTAAATGCGCTGCAACCGTGAACCTGTCAACCTCGGCGC 649  
Qy 181 ATTCAACAGAAATGATACGCGCAACAAAGACTTGTGCGTTACTCCACCGAGCTCGC 240  
Db 650 ATCAACAGAAATGATACGCGCAACAAAGACTTGTGCGTTACTCCACCGAGCTCGC 709  
Qy 241 TTGAGCATAGCGCGCCATCAAAAGCTTGTGTCGCGGCTGGAAGCAACCGT 300  
Db 710 TTGAGCATAGCGCGCCATCAAAAGCTTGTGTCGCGGCTGGAAGCAACCGT 769  
Qy 301 GTGCGTGTAGCATGACGCGCTGAGCTGCTGATTTCGGAAGAAACTCACTGTATGCA 360  
Db 770 GTGCGTGTAGCATGACGCGCTGAGCTGCTGATTTCGGAAGAAACTCACTGTATGCA 829  
Qy 361 CGTTATGCAACTTCAACAGCTCGCGCTGTATATCGACCCCGAATCTGTGCGCAATC 420  
Db 830 CGTTATGCAACTTCAACAGCTCGCGCTGTATATCGACCCCGAATCTGTGCGCAATC 889  
Qy 421 GAATCGGAGGCGCTGACTTTTCAATACCGGTAGCGCGCATTTGGTGGCGGCTG 480  
Db 890 GACATCGTAAAGGGCGGACTCTTTCAATACCGGTAGCGCGCATTTGGTGGCGGCTG 949  
Qy 481 AATTACCAACCCCTGCAAGGACATGATTGCTGTGGACGACAGGCAATTCGCGCTGATG 540  
Db 950 AATTACCAACCCCTGCAAGGACATGATTGCTGTGGACGACAGGCAATTCGCGCTGATG 1009  
Qy 541 ATGAACCAAGGTTACAGACCGCAACCGCGAATGGACAAATACACTCGGTTTCGGGTG 600

Db 1010 ATGAAAAACGGTTACAGCAGCGTAACCGTGAATGGACAAATACCTTCGGGTG 1069  
Qy 601 AGCAACGACCGCTGGATCGCTTCTGCTATTCGCAAGTTCGCGCTCATGACACGAA 660  
Db 1070 AGCAACGACCGCTGGATCGCTTCTGCTATTCGCAAGTTCGCGCTCATGACACGAA 1129  
Qy 661 AGCGCGGCGAGCGTGTATTCGCTAGAGGGTCTGCGAGCGGAGCAATTCCTCGGTG 720  
Db 1130 AGCGCGGCGAGCGTGTATTCGCTAGAGGGTCTGCTAGCGGAGCAATTCCTCGGTG 1189  
Qy 721 TCCTACGCGGTATTCCTGATTCGCTCAAAACAAATACCAACTTCCTTGGGTAAAGATT 780  
Db 1190 TCCTCGCGCGGTATTCCTGATTCGCTCAAAACAAATACCAACTTCCTTGGGTAAAGATT 1249  
Qy 781 GCTTATCAATCAACGACGACCGCATCGCTGCGCCATCTGTTAACGCGCAGCGGCGCAT 840  
Db 1250 GCTTATCAATCAACGACCGCATCGCTGCGCCATCTGTTAACGCGCAGCGGCGCAT 1309  
Qy 841 AATTACAGATTGAAGTCTTATAACCTGACCGCTTCTTCTGCGCGGAGCGCATGAC 900  
Db 1310 AATTACAGATTGAAGTCTTATAACCTGACCGCTTCTTCTTATTCGCGTGAAGTACGAT 1369  
Qy 901 GTAAACAGACGCGCAATGCAACCTCTTTTGAATGAGCGCTGATCAATCAATGCGTG 960  
Db 1370 GTCAACAGACGCGGTAAACCAACCTCTTTTACGAATGAGCGCGGAAATCCGACCGGTG 1429  
Qy 961 TCCTCTTTGAAGCGGACTTCGATATACACACCAACCGATGCGCGGCTTACCAACAA 1020  
Db 1430 TCATGTGTAAGCGGATGTCGATTATCAAAACCAAAAGTATCTGCGGTCAACTACAA 1489  
Qy 1021 GGCTCGTTCCTCGA---CGGATTATTCACCTGGACGCGCACTATATACAGAGATTG 1077  
Db 1490 GGTTGTTCCCGATAGAGGATTCTTCACCTTGACAGCTAATACATCAAGAGACTTG 1549  
Qy 1078 GAGAATATATACAAACGACGATGACACCGGATTCGATTCGCTTCTGCTGATCGAC 1137  
Db 1550 GATGAAATCTACAAACGCGATGATGATACCGCTTCAACGCGATTAACCTCGCTTGAC 1609  
Qy 1138 AGCCACCGTTGCAACT---GGGCGGCAACATCGCTTGTGCTTAAACCTTCGCGCAGT 1194  
Db 1610 AGCCATCGTTGCAACTCGGGGGGGGCGACACCGCTGCTGTTTAAACCTTCGCGCAG 1669  
Qy 1195 CGCGTGTGATTGAAACCTTAAACCGCGAGTATTTACTTCAGCGAAGAGATATCCGT 1254  
Db 1670 CGCGTGTGATTGAAACCTTAAACCGCGAGTATTTACTTCAGCGCGCTGTTGTTGCA 1729  
Qy 1255 ACTACGAGCTCGATTCAACACCGCTGAAACCACTAATTTATGGTTTCTCACTGCTGAT 1314  
Db 1730 ACCACGACGATATCCAGCATCCGCTGAAACCAACCACTACGCTTCTCACTGCTGAC 1789  
Qy 1315 CAAATCCAATGGAACGACGCTTTCAGCAGCGGTGAGATATCGTTACGATCATACCAAA 1374  
Db 1790 CAAATCCAATGGAACGACGCTTTCAGTACCGCGGAGGATTCGTTACGATCATACCAAA 1849  
Qy 1375 ATGACGCTCAGGAAATGAAATGCGGAGTGTATGCTTGTGACAAAAACACCGCTGACGC 1434  
Db 1850 ATGACGCTCAGGAAATGAAATGCGGAGTGTATGCTTGTGACAAAAACACCGCTGACGC 1909  
Qy 1435 AATACCTTAAGGCTGAGCGGATTTGCGGTTTGGCGGCGCAACTGAATCAGGCTTGG 1494  
Db 1910 AACACTTATAAAGCTGGAGCGGTTTGTGCGCTTGGCGGCGCAACTGAATCAGGCTTGG 1969  
Qy 1495 CATGTCGGTTTACGACATTAATTCGCGTACCGTGTCCCAATGCGTCCGAAGTGTATTTC 1554  
Db 1970 CGTGTGCTTACGACATTAATTCGCGTACCGTGTCCCAATGCGTCCGAAGTGTATTTC 2029  
Qy 1555 ACTTACAAACCACTGCTGGGTAATTTGGCTGCCCAATCCCACTGAAAGCGGACGCGCAG 1614  
Db 2030 ACTTACAAACCACTGCTGGGTAATTTGGCTGCCCAATCCCACTGAAAGCGGACGCGCAG 2089  
Qy 1615 ACCACCAACCTGCTCTGCAAGCGCGGAGGAAAAAGGTACTTGTGATGCCAAGCTG 1674

QY 61 GTCTTGGCAGCGGATGAAGCTCAACCCGAAACCAACCCGTTTAAAGCAGAGATAAAAGAA 120  
DB 530 GTCTTTGGCGGAGATGAAGCTCAACTGAACCAACCAACCCGTTTAAAGCAGAGATAAAAGCA 589  
QY 121 GTGGCGTTTAAAGACAGCTTAAATGGCTCTGAAACCGTGGAAAGTGTCAACCTCGCGCGC 180  
DB 590 GTGGCGTTTAAAGCGCAGCGCAATGGCTCTGAAACCGTGGTGTGAAACCGCTCAACCTTAAACCGT 649  
QY 181 ATTCAACAGGAATGATAGCGGACCAACAAAGACTTGGTGGCTTACTTCCACGACGCTCGGC 240  
DB 650 ATCAACAAGAATAATGATACGGGACCAACAAAGACTTGGTGGCTTATTTCCACCGATGTGCGC 709  
QY 241 TTGAGCGATAGCGCGCCCATCAAAAGGCTTTGTCTGGCGGGGTGGAAGCGCAACCGT 300  
DB 710 TTGAGCGACAGCGCGCCCATCAAAAGGCTTTGTCTTTCGGCGGTGGAAGCGCAACCGT 769  
QY 301 GTGGGTGTGAGATTTGACGGGTGAGCTTGCCTGATTCGGAAGAAATCTCACTGTATGCA 360  
DB 770 GTGGCGGTGAGCATAGACGGCGTAAACCTGCTGATTTCCGAAGAAATCTGCTGTACGCC 829  
QY 361 CGTTATGGCAACTTCAACAGCTCGCGCTGTCTATCGACCCCGAACTGTGCGGCAACATC 420  
DB 830 CGTTATGGCAACTTCAACAGCTCGCGCTGTCTATCGACCCCGAACTGTGCGGCAACATC 889  
QY 421 GAAATCGCAAGCGCGCTGACTCTTTTCAATACCGGTAGCGGCGCATTTGGGTGGCGCGGTG 480  
DB 890 GACATCGTAAAGGGCGGACTCTTTCAATAACCGGACGGCGCTTGGCGCGCGGTGTG 949  
QY 481 AATTACAAACCTGTGAAGGACATGATTTGCTGTGTTGGACAGCAGGCAATTCGGCGGTGATG 540  
DB 950 AATTACAAACCTGTGAAGGACATGATTTGCTGTGTTGGACAGCAGGCAATTCGGCGGTGATG 1009  
QY 541 ATGAAGAAACGGTTACAGCAGCGCAACCGCGAATGGACAAATACACTCGGTTTCGGTGTG 600  
DB 1010 ATGAAGAAACGGTTACAGCAGCGCAACCGCGAATGGACAAATACACTCGGTTTCGGTGTG 1069  
QY 601 AGCAACGACCGGTGGATCCGCTTTGCTGTATTCGCAACGTCGCGGTGATGAGACCGGAA 660  
DB 1070 AGCAACGACCGGTGGATCCGCTTTGCTGTATTCGCAACGTCGCGGTGATGAGACCGGAA 1129  
QY 661 AGCGGGGCGAGGTGGCTATCGGTAGAGGTGTGGCAGCGGAGCAATATTCGTTGT 720  
DB 1130 AGCGGGGCGAGGTGGCTATCGGTAGAGGTGTGGCAGCGGAGCAATATTCGTTGT 1189  
QY 721 TCGTACAGCGGTATCCCTGATCCGTCGCAACCAATACCAACTCTTCTGGTGAAGATT 780  
DB 1190 TCTGCGCGGTATTCCTGATCCGTCGCAACCAATACCAACTCTTCTGGTGAAGATT 1249  
QY 781 GCTTATCAAAATCAACGACAGCACCGCATCGGCCCATCGTTTAAAGCGGCGAGCGGCGCAT 840  
DB 1250 GCTTATCAAAATCAACGACAGCACCGCATCGGCCCATCGTTTAAAGCGGCGAGCGGCGCAT 1309  
QY 841 AATTACAGGATGAAGAGTCTTAAACCTGACCGCTTCTTCTGGCGGCAAGCGCGATGAC 900  
DB 1310 AATTACAGGATGAAGAGTCTTAAACCTGACCGCTTCTTCTTATTTGGCGTGAAGCTGACGAT 1369  
QY 901 GTAAACAGACGGCGCAATCGCAACCTCTTTTACGAATGACCGCTGATTTCAAAATTTGGCTG 960  
DB 1370 GTCAACAGACGGCGTAAACCAACCTCTTTTACGAATGACCGCGGATTCACCGCGTTG 1429  
QY 961 TCGTCTTTTGAAGCGGACTTTCGATTTATCAGAACCAACCAAGTGGCGGCTTAAACCAACAA 1020  
DB 1430 TCTATGTTAAAGCGGATTCGATTTATCAAAACCAAGTATCTGCGGTCAACTACAA 1489  
QY 1021 GCGTCTTCCCGA ---CGGATTTATCCACCTGACCGCGCAACTATTAATCAGAAAGATTTG 1077  
DB 1490 GGTTCGTTCCCGATAGAGGATTTCTCCACCTTTGACACGTAACCTACAAATCAAAAGGACTTG 1549  
QY 1078 GAGATATATCAACCGCAGCATGACACCGGATTTCAAGCTTTTACTTTGGTATGGAC 1137  
DB 1550 GATGAATCTTACAAACCGCATGATGGATCCCGCTTCAAAAGCAATTACCCTGCTTTGGAC 1609

QY 1138 AGCCAACCGTTTGAACCT---GGCGGCGCAACATCGCTTGTCTGCTTTAAACTTTTCGCCAGT 1194  
DB 1610 AGCCATCCGTTTGAACCTCGGGGGGGCGACACCGCCTGTCTGCTTTTAAACTTTTCGCCAGC 1669  
QY 1195 CGGCGTGAATTTGAAAACCTTAAACCGCGACGATTTATTACTTTTACGAGAAAGATATCCCGT 1254  
DB 1670 CGCGGTGATTTTGAAGAACTTAAACCGCGACGATTTATTACTTTTACGAGCGCGGTGTGTTGCA 1729  
QY 1255 ACTACAGCTCGATTTCAACACCCCGGTGAAAACCACTAAATATGTTTCTCACTGCTCTGAT 1314  
DB 1730 ACCACAGAGTATCCAGCATCCGTTGAAACCAACCACTAGCGTTTCTCACTGCTGAC 1789  
QY 1315 CAATCAATGAACGACGCTGTTCAGCAGCGGTGAGATATCCGTTTACGATCATACCAA 1374  
DB 1790 CAATTTCAATGAACGACGCTGTTCAGTAGCGCGCAGGTATCCGTTTACGATCATACCAA 1849  
QY 1375 ATGAGCGCTCAGGATTTGAATGCGAGTGTCTGCTTGTGACAAACACCGCTCGAGCC 1434  
DB 1850 ATGAGCGCTCAGGATTTGAATGCGAGTGTCTGCTTGTGACAAACACCGCTCGAGCC 1909  
QY 1435 AATACCTTATAAAGGCTGGAGCGGATTTGTCTGCTTTTGGCGGCGCAACTGAAATCAGGCTTGG 1494  
DB 1910 AACCTTATAAAGGCTGGAGCGGATTTGTCTGCTTTTGGCGGCGCAACTGAAATCAGGCTTGG 1969  
QY 1495 CATGTCGTTACGACATTTACTTTCGGCTACCGGTTCGCCAATGCGTTCGAAAGTGTATTTTC 1554  
DB 1970 CGTGTGCGTTACGACATTTACTTTCGGCTACCGGTTCGCCAATGCGTTCGAAAGTGTATTTTC 2029  
QY 1555 ACTTACACACCGTTTCGGGTAAATTTGGCTGCCCAATCCCACTGAAAGCGGCGGACG 1614  
DB 2030 ACTTACACACCGTTTCGGGTAAATTTGGCTGCCCAATCCCACTGAAAGCGGCGGACG 2089  
QY 1615 ACCACCCACACCTCTCTGCAAGCGCGCAGGAAAAAGGTACTTTTGATGCCAACCTG 1674  
DB 2090 ACCACCCACACCTCTCTGCAAGCGCGCAGGAAAAAGGTACTTTTGATGCCAACCTG 2149  
QY 1675 TATCAAAAACAATTACCGCAACTTTCTGCTGAAGAGCAGAGCTGACCAACGCGCGAT 1734  
DB 2150 TATCAAGCAATTTACCGCAATTTCTGCTGAAGAGCAGAGCTGACCAACGCGCGAT 2209  
QY 1735 GTGCGCTGTACTCAGATGAATTTACTACTAGGTATGTGTAGCAATCTTTATTTCCGAAAA 1794  
DB 2210 GTGCGCTGTACTCAGATGAATTTACTACTAGGTATGTGTAGCAATCTTTATTTCCGAAAA 2269  
QY 1795 CCGGAATGCGAGATCAAAATATTCGATAAGSCCCGAATCCCGTGGTCTTGAGCTGACAGCG 1854  
DB 2270 CTGGAATGCGAGATCAAAATATTCGATAAGSCCCGAATCCCGTGGTATCGAGCTGACGCGC 2329  
QY 1855 CGTCTGAATGTGACAAAAGTAGCGTCTTTTGTCTGAGGGGTGGAATTTGTTGCGCTCG 1914  
DB 2330 CGTCTGAATGTGACAAAAGTAGCGTCTTTTGTCTGAGGGGTGGAATTTGTTGCGCTCG 2389  
QY 1915 CTGCGGTATGCGAAAGCAAACTGTGCGGCGCAACAGCTGCTGCAACAGCGCGCG 1974  
DB 2390 CTGCGGTATGCGAAAGCAAACTGTGCGGCGCAACAGCTGCTGCAACAGCGCGCGTTG 2449  
QY 1975 AAAGTGATTTGCGGTGTGCACTACGAAAGCGCGGAAAAATGCGGTGTGTTCTCCCGC 2034  
DB 2450 AAAGTGATTTGCGGTGTGCACTACGAAAGCGCGGAAAAATGCGGTGTGTTCTCCCGC 2509  
QY 2035 CTGACTTATCTGGGTGCGAAAAAGCGCAACGCGCAATACACCGTTTATGAAACAAAG 2094  
DB 2510 CTGACTTATCTGGGTGCGAAAAAGCGTCAAGACGCGCAATACACCGTTTATGAAACAAAG 2569  
QY 2095 GCGCGGGGTAGCGCTTTTCAGAAAAAGGTAAAGATTTACCGGTGCTGAACAGTCCGCT 2154  
DB 2570 GCGTGGGGTACGCTTTTTCAGAAAAAGGTAAAGATTTACCGGTGCTGAACAGTCCGCT 2629  
QY 2155 TATGTGTTTGTATGTAGCGCTTTTACAAACTGCGGTCTAAAAAAGTCTGACTTGGTGCAGGC 2214  
DB 2630 TATGTGTTTGTATGTAGCGCTTTTACAAACTGCGGTCTAAAAAAGTCTGACTTGGTGCAGGC 2689  
QY 2215 GTATATAATGTGTTCAACCGCAAAATACACCACTTGGGATTTCCCTGCGCGGTTGTATAGC 2274



```

/db_xref="GOA:P57025"
/db_xref="UniProt/Swiss-Prot:P57025"
/translation="MKATIALISAIVIGLACSGPAPAPAAEATPAAPASEAPAAE
AAPADAAAPAGNCAATVESNDNFQKDIQVSKACKBEFTITLKHGTQPKTSMGH
NIVKGTEDMDGIFKDGVAADTVYKPPDARVVAHTKLIGGEBESSLTDPAKLADG
BYKFACTPPGHGALMNGKVTLYD"
complement(3191..3568)
/genes="NMA1733"
/notes="Pfam match to entry PF00127 copper-bind, Copper
binding proteins, plastocyanin/azurin family, score
218.60, E-value 9.1e-62"
complement(3212..3259)
/genes="NMA1733"
/notes="PS00196 Type-1 copper (blue) proteins signature."
complement(4017..4439)
/genes="NMA1734"
complement(4017..4439)
/genes="NMA1734"
/notes="NMA1734, unknown, len: 140aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB84962.1"
/db_xref="GI:7380376"
/translation="NNLEQLGRDALLSGLLKQAGWRRLDAAVKLLPANLHPHFOT
ACIEGRVLVLAANMAASRLKWIAPSVLPOLAGLDASIRSVSLVLPKPEPKPTNT
LHLSXALESFSGSAVKLEKHPLEALANLVRHGA"
4306..4315
/notes="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
complement(4486..4495)
/notes="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
4570..7320
/genes="seca"
4570..7320
/genes="seca"
/notes="NMA1735, seca, preprotein translocase SecA subunit,
len: 916aa; similar to many eg. SW:P10408 (SECA_ECOLI)
preprotein translocase SecA subunit from Escherichia coli
(901 aa) fasta scores; E(): 0, 58.0% identity in 920 aa
overlap. Contains pfam match to entry PF01043
SecA protein, SecA protein, amino terminal region and
Prosite match to PS01312 Protein seca signatures."
/codon_start=1
/transl_table=11
/product="preprotein translocase SecA subunit"
/protein_id="CAB84963.1"
/db_xref="GI:7380377"
/db_xref="GOA:Q9JTK7"
/db_xref="UniProt/TREMBL:Q9JTK7"
/translation="MLTNIARKIFGRNDRLLQYKRVARINALEEQMALSDADLQ
AKTAEFKRLADQGLDGLPEAFVCREASRLTGMRHFDVQLIGGVLHDGKTAEM
RTGEGKTVATLAVLNALAGGVHVTVNDYLSRAGIMEPLNFLGLTVGIISD
MQPFDRONAYADITVGTNNRFGFDYLRNMVTDQYDKVQRELNFAVVDVDSILIDE
Query Match 83.7%; Score 1991; DB 1; Length 326301;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 2143; Conservative 0; Mismatches 230; Indels 4; Gaps 1;
QY 1 ATGAAACCAATTACATGCTTCTATTGCGCGGTGTCGCGAGTATTTTCGGCAATCCG 60
DB 195840 ATGAAACCAATTACATGCTTCTATTGCGCGGTGTCGCGAGTATTTTCGGCAATCCG 195781
QY 61 GTCTTGGCAGCGATGAAGCTCAACCCGAAACACACCCGTTTAAAGCAGAGATAAAGAA 120
DB 195780 GTCTTGGCAGCAGTGAAGCTGCAACTGAAACACACCCGTTTAAAGCAGAGATAAAGCA 195721
QY 121 GTGCGGGTTAAGACACAGCTTAATGCGCTGCAACCGTGAACGTTGTAACCTCGCGCGC 180
DB 195720 GTGCGGGTTAAGAGGTGACGCAATGCGCTGCGGTGTGGAAACGCGTCAACCTTAACCGT 195661

```

```

QY 181 ATTCAACAGGAAATGATACGCGACAAACAAAGACTTTGGTGGTACTTCCACCGACGTGGC 240
DB 195660 ATCAAAAGAAATGATACGCGACAAATAGACTTTGGTGGTACTTCCACCGATTTGGC 195601
QY 241 TTGAGCGATAGCGCGCCCATCAAAAGCGTTTGTGTGCGCGCGTGGAGGCAACCGT 300
DB 195600 TTGAGCGACAGCGCGCGCCATCAAAAGCGTTTGTGTGCGCGCGTGGAGGCAACCGT 195541
QY 301 GTCGGTGTGACATGACGGGTGAGCCTGCTGATTCGGAGAAACATCACTGTATGCA 360
DB 195540 GTCGGCGTGAGCATAGACGGGTGTAACCTGCTGATTCGAAGAAACATCGCTGTACGCC 195481
QY 361 GCTTATGCAACTTCAACAGCTCGCGCTGTCTATTCGACCCCGAACTCGTGGCGCAATC 420
DB 195480 GCTTATGCAACTTCAACAGCTCGCGTTGTCTATCGACCCCGAACTCGTACGCAATTT 195421
QY 421 GAAATCGCGAAGGGCGCTGATCTTTTCAATACCGGTAGCGGCGCATTTGGTGGCGCGGTG 480
DB 195420 GAAATCGTGAAGGGCGCAGACTCTTTTCAATACCGCGAGTGGTGTGATTTGGCGCGGTG 195361
QY 481 AATTACCAAAACCTGCAAGGACATGATTTGCTTTGGAGACGACAGGCAATTCGGCGGTGATG 540
DB 195360 AATTACCAAAACCTGCAAGGCGGTGATTTGCTTTGGAGACGACAGGCAATTCGGCGGTGATG 195301
QY 541 ATGAAAAACGGTTACAGCAGCGCGCAACCGGAAATGCAAAATACACTCGGTTTCGGTGTG 600
DB 195300 ATGAAAAACGGTTACAGCAGCGCGTAACCGTGAATGCAAAATACCTCGGTTTCGGTGTG 195241
QY 601 AGCAACGACCGGTGGATGCGCGCTTTGCTGTATTCGAAACGTCGCGGTGATGACCGGAA 660
DB 195240 AGTAACGACCGGTGGATGCTGCTTTGCTGTATTCGAAACGCGCGGTGATGAAACCGGAA 195181
QY 661 AGCGCGCGCGAGCGGTGGCTATCCGTTAGAGGTGCTGCGAGCGGAGCAATTTATCGTGTG 720
DB 195180 AGCGCGCGCAACCGCGGTATCCGTTAGAGGTGCGGTGAAGAAACGATATCCCGGT 195121
QY 721 TCGTCACGCGGTATCCCTGATCCGTCCAAACAACAATACCACTTCTTGGTGAAGATT 780
DB 195120 TCGCGCGCGCATCCCGATCCGTCCAAACAACAATACCACTTCTTGGTGAAGATT 195061
QY 781 GCTTATCAAAATCAACGACGAGCAGCGATCGGCCCATCGTTTAAAGCGCGAGCGGGCAT 840
DB 195060 GCTTATCAAAATCAACGACGAGCAGCGATCGGCCCATCGCTCAACGCGTCAAGCGGGCAT 195001
QY 841 AATTACGATTTGAAGAGTCTTATAACCTGACCGCTTCTTCTGCGCGGAGCCGATGAC 900
DB 195000 AATTACGATTTGAAGAGTCTTATAACCTGACCGCTTCTTCTGCGCGGAGCCGATGAC 194941
QY 901 GTAAACAGACGCGCAATGCCAACCTCTTTTACGAATGAGCGCTGATTTCAAATTTGGCTG 960
DB 194940 GTAAACAGACGCGCAATGCCAACCTCTTTTACGAATGAGTGCCTGATTTCAAATTTGGTTG 194881
QY 961 TCGTCTTTGAAGCGGAGCTTGTGATATCAGAACCAAGTGGCGCGGTAAACACAA 1020
DB 194880 TCGTCTTTGAAGCGGAGCTTGTGATATCAGAACCAAGTGGCGCGGTAAACACAA 194821
QY 1021 GCCTGTTCCGACCGATTTATTCACCTGGACGCGCAACTATAATCAGAAGGATTTGGAG 1080
DB 194820 GCCTGTTCCGACCGATTTATTCACCTGGACGCGCAACTATAATCAGAAGGATTTGGAG 194761
QY 1081 AATATATCAACCGCAGCATGAGACCCCGATTTCAAACGTTTAACTTTTCGGTATGACAGC 1140
DB 194760 GAAATATCAACCGCAGCATGAGACCCCGATTTCAAACGTTTAACTTTTCGGTATGACAGC 194701
QY 1141 CACCGTTGCAACT----GGGCGGCAACATCGCTTGTGCTTAAACTTTTCGCCAGTGG 1196
DB 194700 CATCGTTGCAACTCGGCGGCGGCGGACACCGCTGCTGTTTAAACTTTTCGCCAGTGG 194641
QY 1197 GCGTGAGTTTGAACACTTAAACCGGAGCATTTATTACTTTCAGCGAAAGAGTATCCGCTAC 1256
DB 194640 CCGTGATTTGAACACTTAAACCGGAGCATTTATTACTTTCAGCGCGCGGTGTTGTTGCAAC 194581
QY 1257 TACCAGCTCGATTTCAACACCCCGTGAACCACTAAATTAATGTTTCTCTACTGCTGATCA 1316

```



```

RESULT 4
NMAG22491/c      326301 bp      DNA      linear      BCT 02-SEP-2002
LOCUS
DEFINITION      Neisseria meningitidis serogroup A strain Z2491 complete genome;
segment 6/7.
ACCESSION      AL162757
VERSION        AL162757.2
KEYWORDS       GI:7380371
SOURCE
ORGANISM       Neisseria meningitidis Z2491
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
REFERENCE
1 (bases 1 to 326301)
Parkhill, J., Achtman, M., James, K.D., Bentley, S.D., Churcher, C.,
Klee, S.R., Morelli, G., Basham, D., Brown, D., Chillingworth, T.,
Davies, R.M., Davis, P., Devlin, K., Feltwell, T., Hamlin, N.,
Holroyd, S., Jagers, K., Leather, S., Moule, S., Mungall, K.,
Quail, M.A., Rajandream, M.A., Rutherford, K.M., Simmonds, M.,
Skelton, J., Whitehead, S., Spratt, B.G. and Barrell, B.G.
Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491
Nature 404 (6777), 502-506 (2000)
JOURNAL
MEDLINE        20222556
PUBMED         10761919
REFERENCE      2 (bases 1 to 326301)
Parkhill, J.
AUTHORS        Direct Submission
TITLE          Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
JOURNAL        sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT
Details of N. meningitidis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES
source
1. 326301
/organism="Neisseria meningitidis Z2491"
/mol_type="genomic DNA"
/strain="Z2491"
/db_xref="taxon:122587"
/notes="serogroup: A"
196..311
/notes="REP 2; hmms hit to HMM REP 2 (1 - 109), score:
87.43"
434..453
/notes">= 90% match to ATCCNNNNNNNGGGAAT"
/label=dRS3
498..560
/notes="Correia element; hmms hit to HMM Correia (1 - 62),
score: 78.71"
/label=Correia
562..604
/notes="Correia element; hmms hit to HMM Correia (114 -
156), score: 51.92"
/label=Correia
complement(667..676)
/notes="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
complement(668..1813)
/genes="dapB"
/genes="dapB"
/genes="dapB"
/EC number="3.5.1.18"
/notes="NMA1730, dapB, succinyl-diaminopimelate
desuccinylase, len: 381aa; similar to many eg. SW.P24176
(DAPE ECOI) dapB, succinyl-diaminopimelate desuccinylase
from Escherichia coli (375 aa) fasta scores; E(): 0, 56.9%
identity in 371 aa overlap. Contains Pfam match to entry
PF01546 Peptidase_M20."
/codon_start=1
/transl_table=11
/product="succinyl-diaminopimelate desuccinylase"
/protein_id="CAB84958.1"

```

```

/db_xref="GI:7380372"
/db_xref="GOA:Q9JTL0"
/db_xref="UniProt/TREMBL:Q9JTL0"
/transl_table="MTETOSLELAKALISRPSTVTPDRDCOKLVERLYKIGFAAEEL
HGDTKNWLRGRTKVPVCFAGHTDVVTFGPVEKWDSPPEFTEEDRGLYGAADM
KTSIACFTACERFVAEPDHQGSIALITSDDEGDAIDGTTKVDVVKARGELIDYC
IVGEPTAVDKLGDMIKNGRSGLSGNLTIVKQGHYAYPHLAINPHTFAPALLELTQ
EVMDEGNKYFPPTSFQISNNGTGTATNIPGELNVKFRPFSTESTAGLQKRVHAI
LDKHGVQYDLOWSCSQPQLTQAGKLTQVARAAIAETCGIEAELSTTGTSDDGRFTKA
IAKELIELGPSNATIHQINENVRLLDDIPKLSAVYEGILARLLAGNAV"
complement(836..1786)
/genes="dapB"
/notes="Pfam match to entry PF01546 Peptidase_M20, score
252.50, E-value 5.7e-72"
complement(1863..1872)
/notes="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
complement(1927..2547)
/genes="NMA1731"
complement(1927..2547)
/genes="NMA1731"
/notes="NMA1731, conserved hypothetical protein, len:
206aa; similar to SW:P44013 (Y552_HAEIN) hypothetical
protein from Haemophilus influenzae (207 aa) fasta scores;
E(): 0, 59.6% identity in 203 aa overlap."
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAB84959.1"
/db_xref="GI:7380373"
/db_xref="UniProt/TREMBL:Q9JQ23"
/transl_table="MLTPKSCDLFNIPFFQSQLKKYQESIPQIKADYKENWQ1WQQ
LIQOAAELGADPAPPHIERWCGWQVRAHFAYFYKQVKNSSAAILSLINRRRLSV
SLDWHCYKADVPALPDYNRLDNFTDKYASFDMMHGAESYDYVRTVAQQNESDR
KLQDSDFFCIGKHIERDLGRQDVAKWIAETVEDLLPYEACHGK"
complement(2551..2558)
complement(2570..2579)
/notes="Core DNA uptake sequence: gcgcgtctgaa"
/label=DUS
complement(2581..3090)
/genes="NMA1732"
complement(2581..3090)
/genes="NMA1732"
/notes="NMA1732, conserved hypothetical protein, len:
169aa; similar to SW:P43971 (Y243_HAEIN) hypothetical
protein from Haemophilus influenzae (172 aa) fasta scores;
E(): 0, 49.7% identity in 169 aa overlap."
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAB84960.1"
/db_xref="GI:7380374"
/db_xref="UniProt/TREMBL:Q9JTK9"
/transl_table="MNPFTQSVTFAPRIEMLYACHGKVRFRFCQGVAMLSDYTAENG
NQLVQITRIQTFYNAAPLHDEENFPLLLQYAPQASVDLRFQVSLHGN
WDAAAFKLEADNAYIPDAEAFKRFVAGYDVHLAIEEPLFDMGKMFIPKELITEIG
EIMARRRR"
complement(3185..3736)
/genes="NMA1733"
complement(3185..3736)
/genes="NMA1733"
/notes="NMA1733, H.8 outer membrane protein, len: 183aa;
strongly similar to SW:P07212 (H8_NEIME) H.8 outer
membrane protein from Neisseria meningitidis (183 aa)
fasta scores; E(): 0, 96.7% identity in 183 aa overlap.
Contains Pfam match to entry PF00127 copper-bind_Copper
binding proteins, plastocyanin/azurin family and Prosite
match to PS00196 Type-1 copper (blue) proteins signature."
/codon_start=1
/transl_table=11
/product="H.8 outer membrane protein"
/protein_id="CAB84961.1"
/db_xref="GI:7380375"

```

|    |        |  |        |
|----|--------|--|--------|
| QY | 241    | TTGAGCGTAGCGCGCCATCAAAAAGCTTTGCTGTCGCGCGGTGGAAGCAACCGT       | 300    |
| Db | 242356 | TTGAGCGTAGCGCGCGCCATCAAAAAGCTTTGCTGTTCCGCGCGGTGGAAGCAACCGT   | 242397 |
| QY | 301    | GTCCGTGTGAGCAATTAAGCGCGTGAAGCTTCTGATTTGGAAGAAATCACTGATGCA    | 360    |
| Db | 242296 | GTCCGTGTGAGCAATTAAGCGGTGAAACCTGCTGATTTGGAAGAAATCACTGATGCA    | 242237 |
| QY | 361    | CGTTATGCGCAATTAAGCGCTGCGGCTGTCTATCGACCCGAACTCGTGGCGCAATC     | 420    |
| Db | 242236 | CGTTATGCGCAATTAAGCGCTGCGGTTGTCTATCGACCCGAACTCGTGGCGCAATC     | 242177 |
| QY | 421    | GAAATCGCAAGGCGCTGACTCTTTCAATACCGGTAGCGCGCATTTGGTGGCGCGT      | 480    |
| Db | 242176 | GAAATCGTGAAGGCGCAGACTCTTTCAATACCGGATGGTGGCATTTGGCGCGCGT      | 242117 |
| QY | 481    | AATTACCAACCTGCAAGGACATGATTTGCTGTTGGACGACAGGCAATTCGCGGTGAT    | 540    |
| Db | 242116 | AATTACCAACCTGCAAGCGCTGATTTGCTGTTGGACGACAGGCAATTCGCGGTGAT     | 242057 |
| QY | 541    | ATGAAGACGTTACAGACCGCACCGGCAATGGAACAAATACACTCGTTTCGGTGTG      | 600    |
| Db | 242056 | ATGAAGACGTTACAGACCGCTGCTGTTGCTGATTTGCAACGTCGCGGTGATGCTG      | 241997 |
| QY | 601    | AGCAACGACCGGTGATGCGCTTTGCTGATTTGCAACGTCGCGGTGATGCTG          | 660    |
| Db | 241996 | AGTAACGACCGGTGATGCTGTTGCTGATTTGCAACGTCGCGGTGATGCTG           | 241937 |
| QY | 661    | AGCGCGGCGAGCGGTGCTATCCGTGAGGTGCTGCGAGCGGCAATTAATTCGTTGT      | 720    |
| Db | 241936 | AGTGGCGGAAACCGAGGCTATGCTGTTGAAGGGAAGGCAATTAATTCGTTGT         | 241877 |
| QY | 721    | TCGTACCGGCTATCCCTGATCGTCCAAACCAATACCAAACTTTCTGCGTGAAT        | 780    |
| Db | 241876 | TCGGCACCGGTATCCCTGATCGTCCAAACCAATACCAAACTTTCTGCGTGAAT        | 241817 |
| QY | 781    | GCTTATCAATCAACGACGACCGCATCGGCCCATCGTTTAAAGCGGACGAGGCAAT      | 840    |
| Db | 241816 | GCTTACCAATTAACGATTAACACCGCATCGGCCCATCGTTTAAAGCGGACGAGGCAAT   | 241757 |
| QY | 841    | AATTACAGATTAAGAGCTTTATTAACCTGACCGCTTTCTGCGCGGAGACCGCATGAC    | 900    |
| Db | 241756 | AATTACAGATTAAGAGCTTTATTAACCTGACCGCTTTCTGCGCGGAGACCGCATGAC    | 241697 |
| QY | 901    | GTAAACAGACGCGCAATGCGAACCTTTTACGATGACCGCTGATTAATTTGGCTG       | 960    |
| Db | 241696 | GTAAACAGACGCGCAATGCGAACCTTTTACGATGATGATGCTGATTAATTTGGCTG     | 241637 |
| QY | 961    | TCGCTCTTTGAAGCGGACCTTCGATTAACAGCAACCAAGTGGCGCGGTAAACACAA     | 1020   |
| Db | 241636 | TCGCTCTTTGAAGCGGACCTTCGATTAACAGCAACCAAGTGGCGCGGTAAACACAA     | 241577 |
| QY | 1021   | GGCTGTTCCCGACGATTAATTCACCTGGAAGCGCACTATTAATCAGAGATTTGAG      | 1080   |
| Db | 241576 | GGCTGTTCCCGATTAATTCACCTGGAAGCGCACTATTAATCAGAGATTTGAG         | 241517 |
| QY | 1081   | AATATATACACCGCAGCATGACACCGATTTCAAGCTTTTACCTTTCGATGACACAGC    | 1140   |
| Db | 241516 | GAAATATACACCGCAGCATGACACCGATTTCAAGCTTTTACCTTTCGATGACACAGC    | 241457 |
| QY | 1141   | CAACCGTTGCAACT---GGCGGCGCAACATCGCTGCTGCTTTAAACCTTTCCGCACTCG  | 1197   |
| Db | 241456 | CATCCGTTGCAACTCGGCGGCGGCGACACCGCTGCTGCTTTAAACCTTTCCGCACTCG   | 241397 |
| QY | 1198   | CGTGAAGTTGAAAACCTTAACCGCAGCATTAATTAATTCAGCGAAAGATATCCCGTACT  | 1257   |
| Db | 241396 | CGTGAAGTTGAAAACCTTAACCGCAGCATTAATTAATTCAGCGGCGCTGTTGTTGGAAC  | 241337 |
| QY | 1258   | ACCAGCTCGATTCACACCCCGTGAACCACTTAATTAATTTGCTTCACTGCTGATCA     | 1317   |
| Db | 241336 | ACCAGCATTCAGCATCCGTTGAACCACTTAATTAATTTGCTTCACTGCTGATCA       | 241277 |
| QY | 1318   | ATCCAATGAACGAGCTGTTCAGACGCGTGCAGATATCCGTTACGATCATACCAAAATG   | 1377   |
| Db | 241276 | ATCCAATGAACGAGCTGTTCAGTAGCGCGAGATATCCGTTACGATCATACCAAAATG    | 241217 |
| QY | 1378   | ACGCTCAGGAATTAATGCGCGAGTGTATGCTTGTGCAAAAACACCGCTGACCCAAAT    | 1437   |
| Db | 241216 | ACGCTCAGGAATTAATGCGCGAGTGTATGCTTGTGCAAAAACACCGCTGACCCAAAT    | 241157 |
| QY | 1438   | ACTTATAAGGCTGAGCGGATTTGTCGTTTGGCGCGCACTGAATCAGGCTTTGGAT      | 1497   |
| Db | 241156 | ACTTATAAGGCTGAGCGGTTTGTGTCGCTTTGGCGCGCACTGAATCAGGCTTTGGAT    | 241097 |
| QY | 1498   | GTCCGTACGACATTAATTCGCGGTACCGTGTCCCAATGCGTCCGAAAGTGTATTTCACT  | 1557   |
| Db | 241096 | GTCCGTACGACATTAATTCGCGGTACCGTGTCCCAATGCGTCCGAAAGTGTATTTCACT  | 241037 |
| QY | 1558   | TACAAACGAGTTTCGGGTAAATTTGGTGCCTAATCCCACTGAAAGCGGAGCGGACGAC   | 1617   |
| Db | 241036 | TACAAACGAGTTTCGGGTAAATTTGGTGCCTAATCCCACTGAAAGCGGAGCGGACGAC   | 240977 |
| QY | 1618   | ACCCACACCTGCTCTCAAGCGCGCAGCAAAAAGGTACTTTTGGATGCCAACCTGTAT    | 1677   |
| Db | 240976 | ACCCACACCTGCTCTCAAGCGCGCAGCAAAAAGGTACTTTTGGATGCCAACCTGTAT    | 240917 |
| QY | 1678   | CAAAAATAATACCGCAACTTCTGCTGAAGACGCAAGAGCTGACCAACGCGCGATGTC    | 1737   |
| Db | 240916 | CAAAAATAATACCGCAACTTCTGCTGAAGACGCAAGAGCTGACCAACGCGCGACTCCC   | 240857 |
| QY | 1738   | GGCTGCTCAGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 1797   |
| Db | 240856 | GGCTGCTCAGATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 240797 |
| QY | 1798   | GAAATGCGAGATGCAAAATATCGATAAGCCCGCAATCCGTTGCTGAGCTGACAGCGCGT  | 1857   |
| Db | 240796 | GAAATGCGAGATGCAAAATATCGATAAGCCCGCAATCCGTTGCTGAGCTGACAGCGCGT  | 240737 |
| QY | 1858   | CTGAATGTGACAAAAGTAGCGCTTTTGTCTGAGGCTGCAAAATTTGTCGCGCTCGCTG   | 1917   |
| Db | 240736 | CTGAATGTGACAAAAGTAGCGCTTTTGTCTGAGGCTGCAAAATTTGTCGCGCTCGCTG   | 240677 |
| QY | 1918   | GGTTATGCGAAAAGCAAACTGTGCGGCGCAACACAGCTGCTGTCACACAGCGCGCGAAA  | 1977   |
| Db | 240676 | GGTTATGCGAAAAGCAAACTGTGCGGCGCAACACAGCTGCTGTCACACAGCGCGCGAAA  | 240617 |
| QY | 1978   | GTGATTCGCGGTGTCGACTACGAAAGCCGAGCGCAAAAATGCGGTGCTGCTCCCGCGCTG | 2037   |
| Db | 240616 | GTGATTCGCGGTGTCGACTACGAAAGCCGAGCGCAAAAATGCGGTGCTGCTCCCGCGCTG | 240557 |
| QY | 2038   | ACTTATCTGGGTGCGAAAAGCCCAAGACGCGCAATACACCGTTTATGAAAACAAAGGCG  | 2097   |
| Db | 240556 | ACTTATCTGGGTGCGAAAAGGTCAAAGACGCGCAATACACCGTTTATGAAAACAAAGGCG | 240497 |
| QY | 2098   | CGGGTACGCTTTGCGAAAAGGTAAGATTAACCGCTGCGCTGCAACAGCTCGGCTTAT    | 2157   |
| Db | 240496 | CGGGTACGCTTTGCGAAAAGGTAAGATTAACCGCTGCGCTGCAACAGCTCGGCTTAT    | 240437 |
| QY | 2158   | GTGTTTGTATGTACGCTTCTCAAACTGCTAAAAACCTGACTTTGCGTGCAGCGCGTA    | 2217   |
| Db | 240436 | GTGTTTGTATGTACGCTTCTCAAACTGCTAAAAACCTGACTTTGCGTGCAGCGCGTA    | 240377 |
| QY | 2218   | TATAATGTGTTCAACCGCAATATACCACTTGGGATTCCTGCGCGGTTTGTATAGCTAC   | 2277   |
| Db | 240376 | TATAATGTGTTCAACCGCAATATACCACTTGGGATTCCTGCGCGGTTTGTATAGCTAC   | 240317 |
| QY | 2278   | AGCACCAACGCGGTGCGAGATGCGAAGGCTTAGCCGCTAGCGGCTCAGCGCTCAGCG    | 2337   |
| Db | 240316 | AGCACCAACGCGGTGCGAGATGCGAAGGCTTAGCCGCTAGCGGCTCAGCGGCTCAGCG   | 240257 |
| QY | 2338   | CGTAATTAACGCGGTATCGCTGATTTGGAAGTTTGA 2373                    |        |
| Db | 240256 | CGTAATTAACGCGGTATCGCTGATTTGGAAGTTTGA 240221                  |        |

|    |      |   |      |
|----|------|---|------|
| Db | 7149 | TCGCTCTTTGAAGCGCGGACTTCGATTATACAGAAAAACCAAGTGGCGCGGTTAAACACAAA  | 7099 |
| Qy | 1021 | GGCTCGTTCCCGACGCGATTATCCACCTGGAAGCGCAACTATATCAGAAGGATTTGGAG     | 1080 |
| Db | 7089 | GGCTCGTTCCCGATGGATTATCCACTGGAAGCGCAACTATATCAGAAGGATTTGGAC       | 7030 |
| Qy | 1081 | AATATATACAAACCGCAGCATGGACACCGGATTTCAAACGTTTACTTTGCGGTATGGACAGC  | 1140 |
| Db | 7029 | GAATATATACAAACCGCAGCATGGACACCGGATTTCAAACGTTTACTTTGCGGTATGGACAGC | 6970 |
| Qy | 1141 | CAACCGTTGGCAACT--GGGCGGCGCAATCGCTTGTGCTTTAAAACTTTGCGCAGTCGG     | 1197 |
| Db | 6969 | CATCGTTGCAACTCGGGGGGGCGGACCGGCTGTGCTTTAAAACTTTGCTGACGCCG        | 6910 |
| Qy | 1198 | CGTGAGTTTGAATACTTAAACCGCGACGATTATTACTTCAGCGGAAAGATATCCCGTACT    | 1257 |
| Db | 6909 | CGTGATTTTGAATACTTAAACCGCGACGATTATTACTTCAGCGGCGCGTGTGTTGTCGAACC  | 6850 |
| Qy | 1258 | ACGAGCTCGATTCAACACCCCGTGAAACACACTAATTATGGTTTCTCAGTGTCTGATCAA    | 1317 |
| Db | 6849 | ACGAGCAGTATCCAGAGTATCCGGTGAACACCAACTACGGTTTCTCAGTGTCTGACCAA     | 6790 |
| Qy | 1318 | ATCCAATGGAAACGACGCTGTTTCAGCAGCGCGTGACATATCCGTTACGATCATACCAAAATG | 1377 |
| Db | 6789 | ATTCAATGGAAACGACGCTGTTTCAGTAGCGCGCAGGTATCCGTTACGACCAACACCAAAATG | 6730 |
| Qy | 1378 | ACGCTCATGGAATTAATGCGGAGTGTATGCTTGTGACAAAAACACCGCTGACGCCAAT      | 1437 |
| Db | 6729 | ACGCTCATGGAATTAATGCGGAGTGTATGCTTGTGACAAAAACACCACTGACGCCAAC      | 6670 |
| Qy | 1438 | ACTTATAAGGCTCGAGCGGATTTGTGCGTTTGGCGCGCAACTGAATCAGGCTTGCGCAT     | 1497 |
| Db | 6669 | ACTTATAAGGCTCGAGCGGTTTGTGCGTTTGGCGCGCAACTGAATCAGGCTTGCGCT       | 6610 |
| Qy | 1498 | GTCGGTTACGACATTACTTCCGGCTACCGTGTCCCCAATCGTCCGAAGTGTAFTTCACT     | 1557 |
| Db | 6609 | GTCGGTTACGACATTACTTCCGGCTACCGTGTCCCCAATCGTCCGAAGTGTAFTTCACT     | 6550 |
| Qy | 1558 | TACAAACACGGTTCCGGTAAATTGCTGCCCAATCCCACTGAAAGCCGAGCGCAGCACC      | 1617 |
| Db | 6549 | TACAAACACGGTTCCGGTAAATTGCTGCCCAATCCCACTGAAAGCCGAGCGCAGCACC      | 6490 |
| Qy | 1618 | ACCACACCCCTGTCTCTGCAAGCGCGACGAAAAAGGTACTTTGGATGCCAACCCTGTAT     | 1677 |
| Db | 6489 | ACCACACCCCTGTCTCTGCAAGCGCGCAGCGAAAAAGGTACTTTGGATGCCAACCCTGTAT   | 6430 |
| Qy | 1678 | CAAAAACAAATTACCGCACTTTGCTGAAGACGACGAAGCTGACCAACGAGGGCGATGTC     | 1737 |
| Db | 6429 | CAAAAGCAATTAACCGCAATTCCTGTCTGAAGACGACGAAGCTGACCAACGAGGGCACTCCC  | 6370 |
| Qy | 1738 | GGCTGTACTCAGATGAATTACTACTAGCGTATGTGTAGCAATCTTATTCGAAAAACCG      | 1797 |
| Db | 6369 | GGCTGTACTGAGGAATAATGCTTACTACGTTATATGACGACGCCCTCACAAGAAAAACTG    | 6310 |
| Qy | 1798 | GAATGGCGNATGCAAAATATCGATAAGCCCGGAATCGTGGTCTTGAGCTGACAGGCCGT     | 1857 |
| Db | 6309 | GATTTGGCAGATGAATAATATCGAAGGCCCAAGATCCGCGGTATTCGAGCTGACAGGCCGT   | 6250 |
| Qy | 1858 | CTGAATGTGACAAAAGTAGCGTCTTTTGTCTCAGGCGCTGGAATTTGTTCCGCTCGCTG     | 1917 |
| Db | 6249 | CTGAATGTGACAAAAGTAGCGTCTTTTGTCTCAGGCGGTGGAATTTGTTCCGCTCGCTG     | 6190 |
| Qy | 1918 | GGTTATGCAAAAGCAAACTGTGCGGCGACAAACGCTGTGTGTCCAACAGCCGCCGAAA      | 1977 |
| Db | 6189 | GGTTATGCAAAAGCAAACTGTGCGGCGACAAACGCTGTGTGTCCAACAGCCGCCGAAA      | 6130 |
| Qy | 1978 | GTGATTCGCGGTGTGCACTACGAAGCCCGACGCAAAAATGGGGGTGTCTTCCCGCTG       | 2037 |
| Db | 6129 | GTGATTCGCGGTATCGACTATGAAGTCCGAGCGAAAAATGGGCGGTATTTCTCCGCTG      | 6070 |
| Qy | 2038 | ACTTATCTGGGTGCGAAAAAGGCCCAAGACGCGCAATACACCGTTTATGAAAAACAGGGC    | 2097 |
| Db | 6069 | ACCTATCTAGCGCGGAAAAAGGTCAAGACGCGCAATACACCGTTTATGAAAAACAGGGC     | 6010 |

|  |      |  |      |
|--|------|--|------|
| QY   | 2098 | CGGGTACGCTTTTCAGAAAAAGGTAAAGATTACCGTGCTGAACAAGTCGGCTTAT        | 2157 |
| Dd   | 6009 | TGGGGTACGCTTTTCAGAAAAAGGTAAAGATTACCGTGCTGAACAAGTCGGCTTAT       | 5950 |
| QY   | 2158 | GTGTTTGATATGTACGGCTTCTAACAACCTGGCTAAAACCTGCACCTTGGGTGCAGCGCGTA | 2217 |
| Dd   | 5949 | GTGTTTGATATGTACGGCTTCTAACAACCGGCTAAAACCTGCACCTTGGGTGCAGCGCGTG  | 5890 |
| QY   | 2218 | TATAATGTGTTCAACCGCAAATACACCACCTTGGGATTCCTGCGGGTTGTGTATAGCTAC   | 2277 |
| Dd   | 5899 | TACAACCTGTTCAACCGCAAATACACCACCTTGGGATTCCTGCGGGTTGTATAGCTAC     | 5830 |
| QY   | 2278 | AGCACCAACACCGCGGTGCACCGAGATGCCAAAGCCTTAGACCGCTACGCGCCCTCAGCG   | 2337 |
| Dd   | 5829 | AGCACCAACATGCGGTGCACCGCGATGCCAAAGCCTTAGACCGCTACGCGCCCGCAGCG    | 5770 |
| QY   | 2338 | CGTAATTACGCGGTATCGTGGATTGGAAAGTTTTGA                           | 2373 |
| Dd   | 5769 | CGCAATTACGCGGTATCGTGGATTGGAAAGTTTTAA                           | 5734 |
| <br>RESULT 3   |      |  |      |
| AX044033/c   |      |  |      |
| LOCUS  |      |  |      |
| DEFINITION   |      |  |      |
| ACCESSION  |      |  |      |
| VERSION  |      |  |      |
| KEYWORDS   |      |  |      |
| SOURCE   |      |  |      |
| ORGANISM   |      |  |      |
| Neisseria meningitidis   |      |  |      |
| Neisseria meningitidis   |      |  |      |
| Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  |      |  |      |
| Neisseriaceae; Neisseria.  |      |  |      |
| 1  |      |  |      |
| REFERENCE  |      |  |      |
| AUTHORS  |      |  |      |
| Piazza,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C., Maignani,V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M., Scarlato,V., Rappuoli,R., Frazer,C.M. and Grandi,G.  |      |  |      |
| Neisseria genomic sequences and methods of their use   |      |  |      |
| Patent: WO 006791-A 112 09-NOV-2000;   |      |  |      |
| CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)  |      |  |      |
| FEATURES   |      |  |      |
| source   |      |  |      |
| Location/Qualifiers  |      |  |      |
| 1..349980  |      |  |      |
| /organism="Neisseria meningitidis"   |      |  |      |
| /mol_type="unassigned DNA"   |      |  |      |
| /db_xref="taxon:487"   |      |  |      |
| /note="sequence too long, cut in 8 pieces.-seq 1: 1 to 349980 349980 bases-seq 108: 300001 to 649980 349980 bases-seq 109: 60001 to 949980 349980 bases-seq 110: 900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980 349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq 113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to 2272325 172325 bases" |      |  |      |
| <br>ORIGIN   |      |  |      |

|    | Query Match           | 85.7%;   | Score 2038.4;      | DB 6;     | Length 349980; |
|----|-----------------------|--|--------------------|-----------|----------------|
|    | Best Local Similarity | 91.4%;   | Prod. No. 0;       |           |                |
|    | Matches 2172;         | Conservative   | 0; Mismatches 201; | Indels 3; | Gaps 1;        |
| QY | 1                     | ATGAACCAATTACACATGCTTCCTATGCGCGCTGGTCGGCAGTATTTTCGGCAATCGG         | 60                 |           |                |
| Db | 242596                | ATGAACCAATTACAAATGCTCCCTATCGCGCGCTGGTCGGCAGTATTTTCGGCAATCGG        |                    |           |                |
| QY | 61                    | GTCTTTGGCAGCGGATGAAGCTGTCAACCGAAACCAACCCGCTTTAAAGCAGAGATATAAGAA    | 120                |           |                |
| Db | 242536                | GTCTTTGGCAGCAGATGAAGCTGTCAACTGAAACCAACACCCGCTTTAAGCGCAGAGATAAAGCA  | 242477             |           |                |
| QY | 121                   | GTGCGCGTTTAAAGACCAAGCTTAATCGCTGCGCTGCAACCGTGGAAACGTTGTAACCTCGGCCGC | 180                |           |                |
| Db | 242476                | GTGCGCGTTTAAAGTTCAGCGCAATCGCTCGCGCTGTGGAACGGCTCAACCTTAACCGT        | 242417             |           |                |
| QY | 181                   | ATTCAACAGGAATATGATACGGCACAACAAAGACTTTGGTGGCGTTATCTCCACCGACGTGGCG   | 240                |           |                |
| Db | 242416                | ATCAAAACAAGAAATATGATACGGCACAATAAAGACTTTGGTGGCGTTATTCACCGATGTGGCG   | 242357             |           |                |

ELEGANRGTSGFLERHOTDYQNYLTGHSTAKOSQYVGHVTEIDENWATVEVKNR  
FAVSDSLEIHPHSGNQTIKLEQWTRKGQVDVAPGNGIQVKIPNMQGKEKALIVLN  
p"

gene  
complement (4380..4766)  
/gene="NMB1665"  
CDS  
complement (4380..4766)  
/gene="NMB1665"  
/note="conserved hypothetical protein; identified by  
Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAF42014.1"  
/db\_xref="GI:7226918"  
/translation="MKILLDTHALLWLLDDKLGISARKLIENPRNAIFVSAASIW  
EISIKONKGLLPIPEFFDLQDEPEMLPIGLPHAKQAGSLPEIHKDPDFDRLIAQT  
QAGFELMTVDYRIFQYIGIRVNASS"  
complement (4763..5011)  
/gene="NMB1666"  
CDS  
complement (4763..5011)  
/gene="NMB1666"  
/note="hypothetical protein; identified by Glimmer2;  
putative"

gene  
complement (4380..4766)  
/gene="NMB1665"  
CDS  
complement (4380..4766)  
/gene="NMB1665"  
/note="conserved hypothetical protein; identified by  
Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAF42015.1"  
/db\_xref="GI:7226919"  
/translation="MFQANIHOAKNLSOLIORABAGEIIVIIKAGKPCVQLIGIEKP  
ARNARLKFMSHMENTDRIEEDNETAALFEESAL"  
complement (5368..5574)  
/gene="NMB1667"  
CDS  
complement (5368..5574)  
/gene="NMB1667"  
/note="hypothetical protein; identified by Glimmer2;  
putative"

gene  
complement (4380..4766)  
/gene="NMB1665"  
CDS  
complement (4380..4766)  
/gene="NMB1665"  
/note="conserved hypothetical protein; identified by  
Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAF42016.1"  
/db\_xref="GI:7226920"  
/translation="MPSEKFRHFELVADMLTDDVGAHDHADVRACLTVVLSALQRL  
LVEVAVEAGNAASARVADAALCL"  
complement (5734..8109)  
/gene="NMB1668"  
CDS  
complement (5734..8109)  
/gene="NMB1668"  
/note="similar to PID:1666679 percent identity: 97.98;  
identified by sequence similarity; putative"

gene  
complement (4380..4766)  
/gene="NMB1665"  
CDS  
complement (4380..4766)  
/gene="NMB1665"  
/note="conserved hypothetical protein; identified by  
Glimmer2; putative"  
/codon\_start=1  
/transl\_table=11  
/product="hemoglobin receptor"  
/protein\_id="AAF42017.1"  
/db\_xref="GI:7226921"  
/translation="MKPLQMLPIAALVGSIFGNVPLAADAETTTTPVKAETKAVRVK  
GORNAPAAVERNLNRIKOEIRDNKDLVRYSTDVGLSDSGRHOGEFVARGVEGNRVG  
VSDGNVLPDSNSLYARVGNFSRLSIDPELVNLEIYKVGASFTNGSALGGV  
NTQTLQDRLLDLRQFGVMKNGHSTNRNWTNLGFGVSDNRVDAALLYSQRGHE  
TESAGNRVAVDEGSGANISARGIPDSKHYHSFIGKIAIYQINDNHRIGASLGE  
QQGHNYTVEESNYLTASSWREADDVNRNANLFEYMPDPSNLSLKLADFDYORTKY  
AAVNNKGSPPMDYSTWYRNQKOLDEIYNRSMDFRFRFTLRLLDPSHLQGGHRL  
SFTKTVSRDFDLNARDYVFSRVRVTTSSIQHPVKTNYGFSLSDOIOWNDVFSR  
AGIRVDHVTQWTEFLNAEACHADKTPPAANTYKHSFGVGLAQLNQRVUCYDITSG  
YRVPNASEVYFTINGSGNLPENLKAERSTHTLSLQGRSEKGLDNLVQSNYRN  
FLSEBQKLTSTPGCTENAYIGSDPYEKLQWQMNIDKARIGIELTGRNLND  
KVASFVPEGMKLGSLGSLKSGNSLLSTQPLKVIAGIDYSPSEKGVFSRLTY  
LGAKKVDAQYTVYENKGTPLQKVKQDYPWLKSAVYVDMYGFYKPAKNTLRAGV  
YNLFNRKYTTWDSLRGLYSYTTNADVRDCKGLDRYRPNRYAVSLEWKF"  
complement (8304..8933)  
/gene="NMB1669"  
CDS  
complement (8304..8933)  
/gene="NMB1669"  
/note="similar to GP:3091146 percent identity: 57.92;

identified by sequence similarity; putative"  
/codon\_start=1  
Query Match 85.7%; Score 2038.4; DB 1; Length 14228;  
Best Local Similarity 91.4%; Pred. No. 0;  
Matches 2172; Conservative 0; Mismatches 201; Indels 3; Gaps 1;  
Qy 1 ATGAACCATTAACATGCTTCTATTGCGCGCTGCTGCGCAGTATTTTCGGCAATCG 60  
Db 8109 ATGAACCATTAACATGCTTCTATTGCGCGCTGCTGCGCAGTATTTTCGGCAATCG 8050  
Qy 61 GTCTTTGGCAGCGGATGAAGTGTCAACCGGAAACCAACACCGCTTTAAAGCAGAGATAAAGAA 120  
Db 8049 GTCTTTGGCAGCAGATGAAGTGTCAACCGGAAACCAACACCGCTTTAAAGCAGAGATAAAGCA 7990  
Qy 121 GTGCGCGTTAAAGACCAAGTGTCAACCGCTGCAACCGTGGAAACGTTCAACCTTCGCGCG 180  
Db 7989 GTGCGCGTTAAAGTGTCAACCGCTGCGCTGCGCTGCGCTGCGCTCAACCTTAAACCGT 7930  
Qy 181 ATTCAACAGGAATATACGCGACAAACAAAGACTTTGGTGGCTTACTTCAACCGACGTCGCG 240  
Db 7929 ATCAACAGGAATATACGCGACAAATTAAGACTTTGGTGGCTTACTTCAACCGACGTCGCG 7870  
Qy 241 TTGACGATAGCGCGCCATCAAAAAGCTTTGCTGTCGCGCGCTGGAAGGCAACCGT 300  
Db 7869 TTGACGACAGCGCGCCATCAAAAAGCTTTGCTGTCGCGCGCTGGAAGGCAACCGT 7810  
Qy 301 GTCGCTGTCAGCATTCAGCGGCTGAGCCTGCTGATTCGAAGAAACATCCTGCTATGCA 360  
Db 7809 GTCGCGCTGAGCATAGACGGTGTAAACCTGCTGATTTCTGAAGAAACATCCTGCTATGCA 7750  
Qy 361 CGTTATGCAACTTCAACAGCTGCGCGCTGCTATTCGACCCCGCAACTGTCGCGCAACATC 420  
Db 7749 CGTTATGCAACTTCAACAGCTGCGCGCTTGTCTATTCGACCCCGCAACTGTCGCGCAACATC 7690  
Qy 421 GAAATCGGAAAGCGCTGACTCTTTCAATACCGGTAGCGCGCATTTGGTGGCGCGCTG 480  
Db 7689 GAAATCGTGAAGCGCGAGACTCTTTCAATACCGCGAGTGGTGATTTGGCGCGCGTGTG 7630  
Qy 481 AATTACCAAAACCTTCAAGGACATGATTTGCTGTTGGAGCAGCAGGCAATTCGGCGTGATG 540  
Db 7629 AATTACCAAAACCTTCAAGGCGCTGATTTGCTGTTGGAGCAGCAGGCAATTCGGCGTGATG 7570  
Qy 541 ATGAATAACCGTTACAGCAGCGCAACCGCAATCGCAATAACATACTCTCGTTTCGGTGTG 600  
Db 7569 ATGAATAACCGTTACAGCAGCGCAACCGCAATCGCAATAACATACTCTCGTTTCGGTGTG 7510  
Qy 601 AGCAACGACCGCTGGATGCGCGCTTTGCTGATTCGCAACCGTTCGCGCTCATGAGACCGAA 660  
Db 7509 AGTAACGACCGCTGGATGCTGCTTTGCTGATTCGCAACCGTTCGCGCTCATGAGACCGAA 7450  
Qy 661 AGCGCGCGGAGCGTGGCTATCCGCTAGAGGTGCTGCGAGCGGAGCAATTTCCGTGGT 720  
Db 7449 AGTGGCGGAAACCGAGGCTATGCTGTGAAGGGGAGGAGCTGTCGCGCGCAATATCCGTGGT 7390  
Qy 721 TCGTCAACCGGTATCCCTGATCCGTCGCAAAACACAAATACCAAACTTCTTGGTGAAGATT 780  
Db 7389 TCGGACCGGTATCCCTGATCCGTCGCAAAACACAAATACCAAACTTCTTGGTGAAGATT 7330  
Qy 781 GCTTATCAAAATCAACGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 840  
Db 7329 GCTTACCAAAATTAACGATAACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7270  
Qy 841 AATTACAGATGAAGAGTCTTATTAACCTGACCGCTTCTTCTGCGCGGAGCCAGCAGCAGCAG 900  
Db 7269 AATTACAGGTTGAAGAGTCTTATTAACCTGACCGCTTCTTCTGCGCGGAGCCAGCAGCAG 7210  
Qy 901 GTAAACAGACGCGCAATGCGCAACCTCTTTTACGAATGAGACCGCTGATTTCAAAATGGCTG 960  
Db 7209 GTAAACAGACGCGCAATGCGCAACCTCTTTTACGAATGAGACCGCTGATTTCAAAATGGCTG 7150  
Qy 961 TCGCTTTTGAAGCGGAGCTTCGATTATCAGACAAACCAAGTGGCGCGCTTAAACAAACAA 1020

```

Db      2264 GGGTACGCGCTTTTCAGAAAAAGGTAAGATTACCGTGGCGTGAACAAGTCGGCTTATGT 2323
Qy      2160 GTTTGATGATGACCGCTTCTACAACTCGCTGCTAATAAAGCTGACTTTGGCTGCGAGCGGTATA 2219
Db      2324 GTTTGATGATGACCGCTTCTACAACTCGCTGCTAATAAAGCTGACTTTGGCTGCGAGCGGTATA 2383
Qy      2220 TAATGTGTTCAACCGCGCAAAATACACCATCTTGGATTCCCTGCGCGGTTTGTATAGCTACAG 2279
Db      2384 TAATGTGTTCAACCGCGCAAAATACACCATCTTGGATTCCCTGCGCGGTTTGTATAGCTACAG 2443
Qy      2280 CACCACCAACCGCGTGCAGCGAGATGCGCAAAAGGCTTAGACCGCTACCGCGCGCTCAGGCGG 2339
Db      2444 CACCACCAACCGCGTGCAGCGAGATGCGCAAAAGGCTTAGACCGCTACCGCGCGCTCAGGCGG 2503
Qy      2340 TAATTACGCGCTATCGCTGCGATTGGAAGTTTGAATTCC 2378
Db      2504 TAATTACGCGCTATCGCTGCGATTGGAAGTTTGAATTCC 2542

RESULT 2
LOCUS   AE002517/c
DEFINITION Neisseria meningitidis serogroup B strain MC58 section 159 of 206
of the complete genome.
ACCESSION AE002517 AE002098
VERSION   AE002517.1 GI:7226912
SOURCE   Neisseria meningitidis MC58
ORGANISM Neisseria meningitidis MC58
REFERENCE 1 (bases 1 to 14228)
AUTHORS Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D.,
Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blait,E.,
Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Massignani,V.,
Piazza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
Rappuoli,R. and Venter,J.C.
Complete genome sequence of Neisseria meningitidis serogroup B
strain MC58
JOURNAL Science 287 (5459), 1809-1815 (2000)
MEDLINE 20175755
PUBMED 10710307
REFERENCE 2 (bases 1 to 14228)
AUTHORS Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D.,
Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blait,E.,
Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Massignani,V.,
Piazza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
Rappuoli,R. and Venter,J.C.
Direct Submission
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
Location/Qualifiers
1. 14228
/organism="Neisseria meningitidis MC58"
/mol_type="genomic DNA"
/strain="MC58"
/db_xref="taxon:122586"
/note="serogroup: B"
complement (77. .283)
/gene="NMB1660"
complement (77. .283)
/gene="NMB1660"
/note="similar to GB:L42023 SP:P43740 PID:1008827
PID:1205973 PID:1221901 percent identity: 69.23;
identified by sequence similarity; putative"

Db      2264 GGGTACGCGCTTTTCAGAAAAAGGTAAGATTACCGTGGCGTGAACAAGTCGGCTTATGT 2323
Qy      2160 GTTTGATGATGACCGCTTCTACAACTCGCTGCTAATAAAGCTGACTTTGGCTGCGAGCGGTATA 2219
Db      2324 GTTTGATGATGACCGCTTCTACAACTCGCTGCTAATAAAGCTGACTTTGGCTGCGAGCGGTATA 2383
Qy      2220 TAATGTGTTCAACCGCGCAAAATACACCATCTTGGATTCCCTGCGCGGTTTGTATAGCTACAG 2279
Db      2384 TAATGTGTTCAACCGCGCAAAATACACCATCTTGGATTCCCTGCGCGGTTTGTATAGCTACAG 2443
Qy      2280 CACCACCAACCGCGTGCAGCGAGATGCGCAAAAGGCTTAGACCGCTACCGCGCGCTCAGGCGG 2339
Db      2444 CACCACCAACCGCGTGCAGCGAGATGCGCAAAAGGCTTAGACCGCTACCGCGCGCTCAGGCGG 2503
Qy      2340 TAATTACGCGCTATCGCTGCGATTGGAAGTTTGAATTCC 2378
Db      2504 TAATTACGCGCTATCGCTGCGATTGGAAGTTTGAATTCC 2542

/codon_start=1
/transl_table=11
/product="DNA-directed RNA polymerase, omega subunit"
/protein_id="AAF42009.1"
/db_xref="GI:7226913"
/transl_table="MARITTEDCTGKISNHFDTLTVARRARQLENGNTPLVDDVRNN
KPTVTALREIAAGHIGTELLTRNK"
complement (342. .959)
/gene="NMB1661"
complement (342. .959)
/gene="NMB1661"
/notes="similar to GB:L42023 SP:P44310 PID:1008829
PID:1205974 PID:1221901 percent identity: 69.11;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="guanylate kinase"
/protein_id="AAF42010.1"
/db_xref="GI:7226914"
/transl_table="MSAYRKGNFIISAAGTGTTLVSRLLANHNLGRVSVSHTRP
PREGEANGVHYHFVSKPEPESLIAQAEFLVADVGVNYGTGAEVNALAAGVDVIL
EIDVQNAQVRDIALPEAVGIFILPPSFDVLAARLNGRTSREVIQRRLSKARHEIEQ
SVLFDVFNVDNDLARAEEEDLHVNAACLKRSQQLGFADLLENS"
1120. .1686
/gene="NMB1662"
1120. .1686
/gene="NMB1662"
/notes="similar to GB:M14040 SP:P07672 GB:X04487 PID:145294
PID:145296 percent identity: 69.10; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="adenine phosphoribosyltransferase"
/protein_id="AAF42011.1"
/db_xref="GI:7226915"
/transl_table="MLVHPKAMSVGALADKIRKIENWPQKILFHDITPVLQSAEYFR
LLVDLIVRYMDQKIDIVAGLARGFTIGALAYQLNVGFVPKRGKLPETVTSQY
ALDVEAAVHTDVAKLSRVLVLDLIATGTGTMGLAGLELRKLGGEIVAAAAILEF
TDLQGGKNIRASGAPLFTLLQNEGCMKG"
complement (1732. .2520)
/gene="NMB1663"
complement (1732. .2520)
/gene="NMB1663"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF42012.1"
/db_xref="GI:7226916"
/transl_table="MNPKIVFDDITLYRKYTDTLRPSVKTAAALRGKILTALAT
GRSLATPEKVRDMAEAGNDVAVTNGQFALLHGKTVREVPMDAGLMGRVCAHLGDL
GMDFAFVGGEGIAVSALSCVCRALKHIAIDFFADKDYFSKPKVYQMLVFAENEMPL
WSDIVREGLKTVRHEEAVDLLIPAGASKTDGIRSVVEALGWENADVNAFQDGLNDVE
MLSEVGFVGAMNGEQAAKAYKVCPSVDEDDGVLRGLQDLGI"
2845. .4200
/gene="NMB1664"
2845. .4200
/gene="NMB1664"
/notes="similar to GB:U00096 SP:P76403 PID:1736789
PID:1736801 PID:1788397 percent identity: 73.62;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="protease, putative"
/protein_id="AAF42013.1"
/db_xref="GI:7226917"
/transl_table="MKAPELLIPAGLERMRAAYDYGADAVYAGSPRYSRLARNNEFA
KLDVLEQIKAEHERNKKPFLTVNTLPHNSKLTFFVADMEPLIAMKDALIMADPGLI
MTVREKWPENPIHLSVQANTYNGVFWQNIQVERIILSELSMERIARTQRCQPI
ELVFTHGALCIAYSGRCLLSGFNHRDQNGCTNSCRWDYKHNATEDAGDAQLL
QGFNFEKAEQNQNFEGQKRHPYADKFLIEESNRGEMMPIMEDSHGTYIMNS
KDLRGIEVVEKLAKIGVDSLKVEGRTKSLYVARVAQSYRAKIDDAVAGRPFDFSLIS

```

|                       |      |   |               |            |              |                   |
|-----------------------|------|---|---------------|------------|--------------|-------------------|
| Query Match           |      | 98.1%;  | Score 2331.8; | DB 1;      | Length 2605; |                   |
| Best Local Similarity |      | 99.0%;  | Pred. No. 0;  |            |              |                   |
| Matches 2356;         |      | Conservative  | 0;            | Mismatches | 22;          | Indels 1; Gaps 1; |
| QY                    | 1    | ATGAACCAATTAACATGCTTCTTATTTGCGCGCTGGTTCGGCAGTATTTTCGGCAATCGG    | 60            |            |              |                   |
| DB                    | 164  | ATGAACCAATTAACATGCTTCTTATTTGCGCGCTGGTTCGGCAGTATTTTCGGCAATCGG    | 223           |            |              |                   |
| QY                    | 61   | GTCTTGGCAGCGATGAAGCTGCAACCGGAACCAACACCCGTTTAAAGCAGAGATAAGAA     | 120           |            |              |                   |
| DB                    | 224  | GTCTTGGCAGCGATGAAGCTGCAACCGGAACCAACACCCGTTTAAAGCAGAGATAAGAA     | 283           |            |              |                   |
| QY                    | 121  | GTGCGGTTTAAAGACCAAGCTTAAATGCGCTGCAACCGTGGAAAGCTGTCAACCTCGGCGCG  | 180           |            |              |                   |
| DB                    | 284  | GTGCGGTTTAAAGACCAAGCTTAAATGCGCTGCAACCGTGGAAAGCTGTCAACCTCGGCGCG  | 343           |            |              |                   |
| QY                    | 181  | ATTCAACAGGAATGATACGCGACAAAGAAAGCTTGGTGGCTTACTCCACCGAGCTGGCG     | 240           |            |              |                   |
| DB                    | 344  | ATTCAACAGGAATGATACGCGACAAAGAAAGCTTGGTGGCTTACTCCACCGAGCTGGCG     | 403           |            |              |                   |
| QY                    | 241  | TTGAGCGATAGCGCGCCCATCAAAAGGCTTTGCTGTGCGCGGCTGGAAGGCAACCGT       | 300           |            |              |                   |
| DB                    | 404  | TTGAGCGATAGCGCGCCCATCAAAAGGCTTTGCTGTGCGCGGCTGGAAGGCAACCGT       | 463           |            |              |                   |
| QY                    | 301  | GTGCGGTGTGAGCAATGACCGCGCTGAGCCTGCTGATTTCCGAAAGAAACTCACTGTATGCA  | 360           |            |              |                   |
| DB                    | 464  | GTGCGGTGTGAGCAATGACCGCGCTGAGCCTGCTGATTTCCGAAAGAAACTCACTGTATGCA  | 523           |            |              |                   |
| QY                    | 361  | CGTTATGGCAAATTCAACAGCTGGCGCTGTCTATCGACCCGAACTGTGTCGCAACATC      | 420           |            |              |                   |
| DB                    | 524  | CGTTATGGCAAATTCAACAGCTGGCGCTGTCTATCGACCCGAACTGTGTCGCAACATC      | 583           |            |              |                   |
| QY                    | 421  | GAAATCGCAAGGCGCTGACTCTTTCATACCGGTAGCGGCGCATTTGGTGGCGCGCTG       | 480           |            |              |                   |
| DB                    | 584  | GAAATCGCAAGGCGCTGACTCTTTCATACCGGTAGCGGCGCATTTGGTGGCGCGCTG       | 643           |            |              |                   |
| QY                    | 481  | AATTACCAACCTGCAAGGACATGATTTGCTGTGGACGACAGGCAATTCGGCGTGATG       | 540           |            |              |                   |
| DB                    | 644  | AATTACCAACCTGCAAGGACATGATTTGCTGTGGACGACAGGCAATTCGGCGTGATG       | 703           |            |              |                   |
| QY                    | 541  | ATGAACCAATTAACATGCTTCTTATTTGCGCGCTGGTTCGGCAGTATTTTCGGCAATCGG    | 600           |            |              |                   |
| DB                    | 704  | ATGAACCAATTAACATGCTTCTTATTTGCGCGCTGGTTCGGCAGTATTTTCGGCAATCGG    | 763           |            |              |                   |
| QY                    | 601  | AGCAACGCGCGTGTGATCCGCTTTCGCTGATTTGCAACGCTGCGGTCATGAGACCGAA      | 660           |            |              |                   |
| DB                    | 764  | AGCAACGCGCGTGTGATCCGCTTTCGCTGATTTGCAACGCTGCGGTCATGAGACCGAA      | 823           |            |              |                   |
| QY                    | 661  | AGCGCGGCGAGCGTGGCTATCCGCTAGAGGCTGCTGGCAGCGGAGCAATTTCCGCTGGT     | 720           |            |              |                   |
| DB                    | 824  | AGCGCGGCGAGCGTGGCTATCCGCTAGAGGCTGCTGGCAGCGGAGCAATTTCCGCTGGT     | 883           |            |              |                   |
| QY                    | 721  | TGCTACGCGGTTATCCCTGATCCGCTGCAACCAAAATACCAACTTCTTGGGTAAGATT      | 780           |            |              |                   |
| DB                    | 884  | TGCTACGCGGTTATCCCTGATCCGCTGCAACCAAAATACCAACTTCTTGGGTAAGATT      | 943           |            |              |                   |
| QY                    | 781  | GCTTATCAATCAACGACAGCACCGCATCGGCCCATCGTTTAAAGCGCAGCAGGGGAT       | 840           |            |              |                   |
| DB                    | 944  | GCTTATCAATCAACGACAGCACCGCATCGGCCCATCGTTTAAAGCGCAGCAGGGGAT       | 1003          |            |              |                   |
| QY                    | 841  | AATTACAGCA-TTGAAGAGTCTTATAACCTGACCGCTTCTTCTGGCGGAAAGCGATGA      | 899           |            |              |                   |
| DB                    | 1004 | AATTACAGCA-TTGAAGAGTCTTATAACCTGACCGCTTCTTCTGGCGGAAAGCGATGA      | 1063          |            |              |                   |
| QY                    | 900  | CGTAAACAGACGCGCAATGCGCAACTCTTTTACGAATGGAAGCTGTCAAAATTTGGCT      | 959           |            |              |                   |
| DB                    | 1064 | CGTAAACAGACGCGCAATGCGCAACTCTTTTACGAATGGAAGCTGTCAAAATTTGGCT      | 1123          |            |              |                   |
| QY                    | 960  | GTGCTCTTTGAAGGCGGACTTCGATATTCAGACCAACCAAGTGGCGGCTTAAACAA        | 1019          |            |              |                   |
| DB                    | 1124 | GTGCTCTTTGAAGGCGGACTTCGATATTCAGACCAACCAAGTGGCGGCTTAAACAA        | 1183          |            |              |                   |
| QY                    | 1020 | AGGCTCGTTCCGACGGAATTTTCCACCTGGACGGCAACTATATCAGAAGGATTTTGA       | 1079          |            |              |                   |
| DB                    | 1184 | AGGCTCGTTCCGACGGAATTTTCCACCTGGACGGCAACTATATATCAGAAGGATTTTGA     | 1243          |            |              |                   |
| QY                    | 1080 | GAATATATACAAACCGCAGATGGACACCCGATTTAAAGTCTTTTACTTTTGGTATGACAG    | 1139          |            |              |                   |
| DB                    | 1244 | GAATATATACAAACCGCAGATGGACACCCGATTTAAAGTCTTTTACTTTTGGTATGACAG    | 1303          |            |              |                   |
| QY                    | 1140 | CCAACCGTTGCAACTGGCGCGCAACATCGCTTGTGCTTAAACCTTTTCGCCAGTTCGGCG    | 1199          |            |              |                   |
| DB                    | 1304 | CCAACCGTTGCAACTGGCGCGCAACATCGCTTGTGCTTAAACCTTTTCGCCAGTTCGGCG    | 1363          |            |              |                   |
| QY                    | 1200 | TCAGTTTGAAACCTTTAAACCGCGAGATTTACTTTACGCGAAAGAGTATCCCGTACTAC     | 1259          |            |              |                   |
| DB                    | 1364 | TCAGTTTGAAACCTTTAAACCGCGAGATTTACTTTACGCGAAAGAGTATCCCGTACTAC     | 1423          |            |              |                   |
| QY                    | 1260 | CAGCTCGATTTCAACACCCCGTGAAACCACTAAATATGTTTCTCAGCTGTCTGATCAAT     | 1319          |            |              |                   |
| DB                    | 1424 | CAGCTCGATTTCAACACCCCGTGAAACCACTAAATATGTTTCTCAGCTGTCTGATCAAT     | 1483          |            |              |                   |
| QY                    | 1320 | CCAATGGAAACGAGCTGTTTACGAGCGGTGAGATATCGTTTACGATCATATCCAAATGAC    | 1379          |            |              |                   |
| DB                    | 1484 | CCAATGGAAACGAGCTGTTTACGAGCGGTGAGATATCGTTTACGATCATATCCAAATGAC    | 1543          |            |              |                   |
| QY                    | 1380 | GCCTCAGGAATTTGAATGCGGAGTGTCTATGCTTGTGACAAAAACACCGCTGCGCAATAC    | 1439          |            |              |                   |
| DB                    | 1544 | GCCTCAGGAATTTGAATGCGGAGTGTCTATGCTTGTGACAAAAACACCGCTGCGCAATAC    | 1603          |            |              |                   |
| QY                    | 1440 | TTATAAGGCTGGAGCGGATTTGCGTTTGGCGCGCAACTGAATCAGGCTGGCATGT         | 1499          |            |              |                   |
| DB                    | 1604 | TTATAAGGCTGGAGCGGATTTGCGTTTGGCGCGCAACTGAATCAGGCTGGCATGT         | 1663          |            |              |                   |
| QY                    | 1500 | CGGTTACGACATTAATCTCCGCTACCGGTGCCCAATCCCAACCTGCGGAGTGTATTTCACTTA | 1559          |            |              |                   |
| DB                    | 1664 | CGGTTACGACATTAATCTCCGCTACCGGTGCCCAATCCCAACCTGCGGAGTGTATTTCACTTA | 1723          |            |              |                   |
| QY                    | 1560 | CAACCACTGCTGGGTAAATGGCTGCCCAATCCCAACCTGCGGAGCGGAGCGAGCCAC       | 1619          |            |              |                   |
| DB                    | 1724 | CAACCACTGCTGGGTAAATGGCTGCCCAATCCCAACCTGCGGAGCGGAGCGAGCCAC       | 1783          |            |              |                   |
| QY                    | 1620 | CCACACCTGTCTCTGAAAGCGCGGAGGAAAGGTAATTTGGATGCGCACTGTATCA         | 1679          |            |              |                   |
| DB                    | 1784 | CCACACCTGTCTCTGAAAGCGCGGAGGAAAGGTAATTTGGATGCGCACTGTATCA         | 1843          |            |              |                   |
| QY                    | 1680 | AAACAAATTAACCGCAATTTCTTGTCTGAAAGCAGAGCTGACCAACCGCGGAGTGTGG      | 1739          |            |              |                   |
| DB                    | 1844 | AAACAAATTAACCGCAATTTCTTGTCTGAAAGCAGAGCTGACCAACCGCGGAGTGTGG      | 1903          |            |              |                   |
| QY                    | 1740 | CTGTACTCAGATGAATTAATCTACTAGGTATGTGTAGCAATCTTATTCGAAAAACCGGA     | 1799          |            |              |                   |
| DB                    | 1904 | CTGTACTCAGATGAATTAATCTACTAGGTATGTGTAGCAATCTTATTCGAAAAACCGGA     | 1963          |            |              |                   |
| QY                    | 1800 | ATGGCAGATGCAAAATATGATAGGCGCGAATCCGTGGTCTTGGATGACAGGCGGCTCT      | 1859          |            |              |                   |
| DB                    | 1964 | ATGGCAGATGCAAAATATGATAGGCGCGAATCCGTGGTCTTGGATGACAGGCGGCTCT      | 2023          |            |              |                   |
| QY                    | 1860 | GAATGTGAAAAAGTAGGCTCTTTTCTGAGGCTGGAATTTGTTCTGGCTCGCTGGG         | 1919          |            |              |                   |
| DB                    | 2024 | GAATGTGAAAAAGTAGGCTCTTTTCTGAGGCTGGAATTTGTTCTGGCTCGCTGGG         | 2083          |            |              |                   |
| QY                    | 1920 | TTATGCGAAAAAGCAACTGTTCGGGCGCAACAGCGCTGTCTCCACAGCGCGCGGAAAGT     | 1979          |            |              |                   |
| DB                    | 2084 | TTATGCGAAAAAGCAACTGTTCGGGCGCAACAGCGCTGTCTCCACAGCGCGCGGAAAGT     | 2143          |            |              |                   |
| QY                    | 1980 | GATTCGCGGTGCACTACGAAAGCCGAGCGGAAATGGGGTGTGTTCTCCGCGCTGAC        | 2039          |            |              |                   |
| DB                    | 2144 | GATTCGCGGTGCACTACGAAAGCCGAGCGGAAATGGGGTGTGTTCTCCGCGCTGAC        | 2203          |            |              |                   |
| QY                    | 2040 | TTATCTGGGTGCGAAAAAGGCAAGACCGCAATACACCGTTTATGAAAAACAGGGCGG       | 2099          |            |              |                   |
| DB                    | 2204 | TTATCTGGGTGCGAAAAAGGCAAGACCGCAATACACCGTTTATGAAAAACAGGGCGG       | 2263          |            |              |                   |
| QY                    | 2100 | GGGTACGCTTTTGCAAAAAAGGTAAAGATTTACCGGTGGCTGGAACAAAGTCCGCTTATGT   | 2159          |            |              |                   |

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2005, 16:06:11 ; Search time 10295 Seconds  
(without alignments)  
11192.472 Million cell updates/sec

Title: US-09-665-358-7  
Perfect score: 2378  
Sequence: 1 ATGAACCATACACATGCT.....GATTGGAAGTTTGAAATTC 2378

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_btg.\*

3: gb\_in.\*

4: gb\_cm.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query |        | DB | ID        | Description        |
|------------|--------|-------|--------|----|-----------|--------------------|
|            |        | Match | Length |    |           |                    |
| 1          | 2331.8 | 98.1  | 2605   | 1  | NGU40861  | U40861 Neisseria g |
| C 2        | 2038.4 | 85.7  | 14228  | 1  | AE002517  | AE002517 Neisseria |
| C 3        | 2038.4 | 85.7  | 34980  | 6  | AX044033  | AX044033 Sequence  |
| C 4        | 1991   | 83.7  | 326301 | 1  | NMA622491 | AL162757 Neisseria |
| 5          | 1988.6 | 83.6  | 3318   | 1  | NMU18558  | U18558 Neisseria m |
| 6          | 1988.6 | 83.6  | 3318   | 6  | I85729    | I85729 Sequence 1  |
| 7          | 1987.2 | 83.6  | 2632   | 1  | NMU40859  | U40859 Neisseria m |
| 8          | 1985.4 | 83.5  | 2652   | 1  | AF105339  | AF105339 Neisseria |
| 9          | 1844.4 | 77.6  | 2568   | 1  | NMU40860  | U40860 Neisseria m |
| 10         | 531    | 22.3  | 1500   | 1  | AF133695  | AF133695 Neisseria |
| 11         | 456.4  | 19.2  | 10062  | 1  | AE006111  | AE006111 Pasteurel |
| 12         | 226.4  | 9.5   | 2424   | 1  | AF319529  | AF319529 Neisseria |
| 13         | 109.8  | 4.6   | 3850   | 1  | NMU73112  | U73112 Neisseria m |
| C 14       | 108.6  | 4.6   | 349061 | 1  | NMA222491 | AL162753 Neisseria |
| 15         | 101.4  | 4.3   | 2454   | 1  | AF319528  | AF319528 Neisseria |
| C 16       | 76.4   | 3.2   | 1415   | 1  | AF319531  | AF319531 Neisseria |
| 17         | 75.8   | 3.2   | 3171   | 1  | NMIRA     | G69214 N.meningiti |
| 18         | 75.8   | 3.2   | 5691   | 1  | AF049349  | AF049349 Neisseria |
| 19         | 74     | 3.1   | 3300   | 1  | NGU16260  | U16260 Neisseria g |

|    |      |      |        |      |          |                       |           |
|----|------|------|--------|------|----------|-----------------------|-----------|
| 20 | 74   | 3.1  | 3300   | 6    | AX685922 | Sequence              |           |
| 21 | 73.6 | 3.1  | 3398   | 1    | NMLBPAG  | X79838 N.meningitidis |           |
| C  | 22   | 72   | 3.0    | 9955 | 1        | AE002504              | Neisseria |
| C  | 23   | 69.2 | 2.9    | 558  | 6        | BD15880               | Neisseria |
| C  | 24   | 69.2 | 2.9    | 558  | 6        | CQ754837              | Sequence  |
| C  | 25   | 69.2 | 2.9    | 918  | 6        | BD15881               | Neisseria |
| C  | 26   | 69.2 | 2.9    | 918  | 6        | CQ754838              | Sequence  |
| 27 | 68.4 | 2.9  | 10799  | 1    | AE004559 | Pseudomonas           |           |
| 28 | 68.2 | 2.9  | 303642 | 1    | AE016923 | Chromobacterium       |           |
| 29 | 66.6 | 2.8  | 2739   | 1    | AF124339 | Neisseria             |           |
| 30 | 66.6 | 2.8  | 2748   | 1    | AF124338 | Neisseria             |           |
| C  | 31   | 66   | 12231  | 1    | AE006069 | Pasteurella           |           |
| 32 | 65.4 | 2.8  | 2809   | 6    | A36258   | Sequence 5            |           |
| 33 | 65.4 | 2.8  | 2809   | 6    | AR179569 | Sequence              |           |
| 34 | 65.4 | 2.8  | 5472   | 1    | NMTBP12B | Z15130 N.meningitidis |           |
| 35 | 65   | 2.7  | 2916   | 1    | NGOTBPA  | N96731 Neisseria      |           |
| 36 | 65   | 2.7  | 3035   | 1    | AF241227 | Neisseria             |           |
| 37 | 65   | 2.7  | 3035   | 6    | AX685925 | Sequence              |           |
| 38 | 65   | 2.7  | 3286   | 6    | AR072080 | Sequence              |           |
| 39 | 64.2 | 2.7  | 2748   | 1    | AF268474 | Neisseria             |           |
| 40 | 64.2 | 2.7  | 2748   | 6    | AX259189 | Sequence              |           |
| 41 | 64.2 | 2.7  | 2762   | 1    | NMTBP1B  | X75166 N.meningitidis |           |
| 42 | 64.2 | 2.7  | 11109  | 1    | AE002402 | Neisseria             |           |
| 43 | 64.2 | 2.7  | 349980 | 6    | AX044029 | Sequence              |           |
| 44 | 63.8 | 2.7  | 2955   | 6    | AR083968 | Sequence              |           |
| 45 | 63.8 | 2.7  | 3000   | 6    | AR083967 | Sequence              |           |

ALIGNMENTS

|            |   |   |   |         |        |                 |                 |
|------------|---|---|---|---------|--------|-----------------|-----------------|
| RESULT 1   | NGU40861  | NGU40861  | Neisseria gonorrhoeae hemoglobin receptor (hmbR) gene, complete | 2605 bp | DNA    | linear          | BCT 13-NOV-1996 |
| LOCUS      | NGU40861  | Neisseria gonorrhoeae hemoglobin receptor (hmbR) gene, complete | 2605 bp   | DNA     | linear | BCT 13-NOV-1996 |                 |
| DEFINITION | NGU40861  | Neisseria gonorrhoeae hemoglobin receptor (hmbR) gene, complete | 2605 bp   | DNA     | linear | BCT 13-NOV-1996 |                 |
| ACCESSION  | U40861  | GI:16666676   |   |         |        |                 |                 |
| VERSION    | U40861.1  | GI:16666676   |   |         |        |                 |                 |
| KEYWORDS   | Neisseria gonorrhoeae   |   |   |         |        |                 |                 |
| SOURCE     | Neisseria gonorrhoeae   |   |   |         |        |                 |                 |
| ORGANISM   | Neisseria gonorrhoeae   |   |   |         |        |                 |                 |
| REFERENCE  | 1 (bases 1 to 2605)   |   |   |         |        |                 |                 |
| AUTHORS    | Stojiljkovic, I., Larson, J., Hwa, V., Anic, S. and So, M.  |   |   |         |        |                 |                 |
| TITLE      | Direct Submission   |   |   |         |        |                 |                 |
| JOURNAL    | Submitted (19-NOV-1995) Igor Stojiljkovic, Microbiology & Immunology, OHSU, L220, Portland, OR 97201, USA |   |   |         |        |                 |                 |
| FEATURES   | Location/Qualifiers   |   |   |         |        |                 |                 |
| source     | 1..2605   |   |   |         |        |                 |                 |
|            | /organism="Neisseria gonorrhoeae"   |   |   |         |        |                 |                 |
|            | /mol_type="genomic DNA"   |   |   |         |        |                 |                 |
|            | /strain="WS11"  |   |   |         |        |                 |                 |
|            | /db_xref="taxon:485"  |   |   |         |        |                 |                 |
|            | 128..1018   |   |   |         |        |                 |                 |
|            | /gene="hmbR"  |   |   |         |        |                 |                 |
|            | 128..1018   |   |   |         |        |                 |                 |
|            | /gene="hmbR"  |   |   |         |        |                 |                 |
|            | /notes="HmbR; truncated product due to premature stop codon"  |   |   |         |        |                 |                 |
|            | codon   |   |   |         |        |                 |                 |
|            | /codon_start=1  |   |   |         |        |                 |                 |
|            | /transl_table=11  |   |   |         |        |                 |                 |
|            | /product="hemoglobin receptor"  |   |   |         |        |                 |                 |
|            | /protein_id="AAB18727.1"  |   |   |         |        |                 |                 |
|            | /translation="MDEQITSCVDNKKPLHMLPIALVGSIFGNPVLAADEAATETT  |   |   |         |        |                 |                 |
|            | PVAKIEKVRVKDQANAPATVERVNLGRVQEMIRNKKLVYRSTVGLSDSGRHQKG  |   |   |         |        |                 |                 |
|            | FVARGVGNRVGSDVSLPDSSENLVARYGNFNSRLSIDPELVNRIEAKGADS   |   |   |         |        |                 |                 |
|            | FTGSCALGGVNYQTLOGHDLILLDDQFGVMKNGYSRNRWNTTLGFGNDRVD   |   |   |         |        |                 |                 |
|            | ALLYSQRGHETSAGERGYPEVGAGSGANIGSRGIPDPSPKHVHFLGLIAVQI  |   |   |         |        |                 |                 |
|            | NDKRRIGRPSFNGQQGHNYTI"  |   |   |         |        |                 |                 |

ORIGIN



Query Match 23.9%; Score 1003; DB 2; Length 209;  
 Best Local Similarity 94.7%; Pred. No. 4.2e-56;  
 Matches 198; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKPLHMLPIAALVGSIFGNPVLAADEAATETTPVKAIEKEVRVKDQLNAPATVERVNLGR 60  
 DB 1 MKPLQMLPIAALVGSIFGNPVLAADEAATETTPVKAIEKAVRVKGQRPAAVERVNLNR 60

QY 61 TQEMIRNOKLVRYSTDVGLSDSRHQKGFPAVRGVEGNRVGVSIDGVSLLPDSENSLYA 120  
 DB 61 TQEMIRNOKLVRYSTDVGLSDSRHQKGFPAVRGVEGNRVGVSIDGVSLLPDSENSLYA 120

QY 121 RYGNFNSRLSDPELVNRIETAKGADSFNTGSGALGGGVNYQTQGHDLILLDDRRQFGVM 180  
 DB 121 RYGNFNSRLSDPELVNRIETAKGADSFNTGSGALGGGVNYQTQGHDLILLDDRRQFGVM 180

QY 181 MKNGYSSRNREWTNLFGVSNDRVDAAL 209  
 DB 181 MKNGYSSRNREWTNLFGVSNDRVDAAL 209

## RESULT 9

Q8GB30 PRELIMINARY; PRT; 763 AA.  
 ID Q8GB30; (TREMBLrel. 23, Created)  
 AC Q8GB30; (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Transferrin-binding protein A.  
 OS Name=tbpa;  
 GN Haemophilus somnus.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Histophilus.  
 OX NCBI\_TaxID=731;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=649;  
 RA Ekins A., Bahrami P., Sijercic A., Maret D., Niven D.F.;  
 RT "Haemophilus somnus possesses two systems for acquisition of  
 transferrin-bound iron."  
 RL J. Bacteriol. 186:4407-4411(2004).  
 DR EMBL; AY184230; AAC11733.1; -.  
 DR GO; GO:0019867; C:outer membrane; IEA.  
 DR GO; GO:0004872; E:receptor activity; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR010948; TonB-hemolactins.  
 DR TIGRFAMs; TIGR01786; TonB-hemolactins; 1.  
 DR TIGRFAMs; TIGR01776; TonB-tbp-lbp; 1.  
 SQ SEQUENCE 763 AA; 86642 MW; 9C239A449699EDD3 CRC64;

Query Match 20.4%; Score 857; DB 2; Length 763;  
 Best Local Similarity 30.2%; Pred. No. 5.2e-46;  
 Matches 242; Conservative 151; Mismatches 321; Indels 86; Gaps 28;

QY 20 PVLA-ADEAATETTPVKAIEKEVRVKDQLNAPATVERV-----NLGRTOQEMIRNOKLV 73  
 DB 21 PALSADNSQEKTKPKAB-----ELGEISVVYANVEPIFSSQSEQLRKNMVSNERDLF 75

QY 74 RYSTVDGLSDSRH-QKGFPAVRGVEGNRVGVSIDGVSLLPDSENSLYA-----GNF 125  
 DB 76 RHEVIGVPEGGRSGNGFAIGVKDKRVAVIVDGI--POAESTMTSTNRYSTERHNGNI 133

QY 126 NSSRLSDPELVNRIETAKGADSFNTGSGALGGGVNYQTQGHDLILLDDRRQFGVMKNGY 185  
 DB 134 NNT-----EYENISSVKQKANSVMNGSGALGGVSTTKETEDFVPEDRTEGFLSKTGY 189

QY 186 SSRNREWTNLFGVSNDRVDAALYSQRCHETESAGERGYPVEGSGAIIRSSRGI 245  
 DB 190 TSKNREWRQVGGGKTRDFFGFAQLTRKWHGTINNGK-----GADITGFGRGK 239

QY 246 PDPKHKYHNFELGKIAYQINDKHRIGPSFNGQOQHNYTIESSYNLTASSWREADDVNRRR 305  
 DB 240 PNPISYHTTSLTKIGYDINNTHFTLPLEDAEKKFTTEKTLGLSDEN-REATDRTPYR 298

QY 306 NANLFYEWTPDSNWSLKLADFYQ-TTKVAANNNKGSFPDPTSTWTWYNNQKLENIYN 364  
 DB 299 RYGLERYN-GLSWLETAKLFAVKQKIEQLSALRGISSY-----NNNLFQLAS-YQ 347

QY 365 RSMTRPKRFTLRWDSPQLQ-LGGHRLSLATFASREFENLRDDVYFSEVSRTTSSI 423  
 DB 348 YIQDTITRGL--STYPLALWRTHRLSSKVEPRDQFLENSQRYFVSLMKKNRPNY 405

QY 424 QH--PVKTYNTGFSLSQIQWNVFSSRADIRYDHTKMTPOELNAEACHADKTPPAAN-T 480  
 DB 406 TEVVPKSRIFSLSLDEVALTDILKATIGGRYDRYVAPQNVTKTLGAKINPFPSKLS 465

QY 481 YKWSGFGVGLAAQNAQWHGYDITSGYRVNASEVYTYNHGSGN--WLPNPNLKAERS 538  
 DB 466 HLSWQ--LGLEYQVHPHQIGYRLSTGPRVPAEDLYFVSRSESTDIETIPNPFLOPETA 523

QY 539 THTLSLQGRSEKGTLDANLYQNNYRNFSLSEQLTTSQDVGCTQMYYYGMCSPYSEK 598  
 DB 524 LNHETLYRFQNYAHFVGLFRTHYHFIQER-----CKNRFWDPPDKSYGAK 575

QY 599 PEW----QMNIDKARIRGLTGLNVTKVASFVPEG--WKLFGSLGYAKSLSGDMSL 652  
 DB 576 CTHSLAQFVNEPKAVIKGIEVSGAVNGS--AFGLSDGLTFLKGS--YSKQNHGDGDL 631

QY 653 LSTPPKVIAGVDYSESPKGVSRSLYTLGAKAKADQYTYENKGRGTPLQKKVQYDP 712  
 DB 632 KSIQFPTVVGIDYE--TEKWSISLTGRYSAAKAKADAETEY-----THDKKVVQWMP 683

QY 713 WLKNSAYVDFDMYGVKLAKNLTLAGVNVNVRKVTWDSLRLGLYS--YSTNAVDRDGK 770  
 DB 684 FLSPSYFVVDLTGQVNLKSNVILNMGIFNLFRNYTTHWDSAYNLFTRGFTSRSLVNQPS 743

QY 771 GLDRYASGRNYAVSLDWKF 790  
 DB 744 GINRFTAPKRNFAASVEVRF 763

RESULT 10  
 Q934H9 PRELIMINARY; PRT; 759 AA.  
 ID Q934H9  
 AC Q934H9;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Transferrin binding protein A.  
 OS Name=tbpa;  
 GN Haemophilus somnus.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Histophilus.  
 OX NCBI\_TaxID=731;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=9L;  
 RX MEDLINE=21945385; PubMed=11948169;  
 DOI=10.1128/JB.184.9.2539-2542.2002;  
 RA Ekins A., Niven D.F.;  
 RT "Identification of fur and fldA homologs and a Pasteurella multocida  
 tbpa homolog in Histophilus ovis and effects of iron availability on  
 their transcription."  
 RL J. Bacteriol. 184:2539-2542(2002).  
 DR EMBL; AY040784; AAK68653.1; -.  
 DR GO; GO:0019867; C:outer membrane; IEA.  
 DR GO; GO:0004872; E:receptor activity; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR010948; Lcfr/trsfrecept.  
 DR InterPro; IPR010949; TonB-hemolactins.  
 DR TIGRFAMs; TIGR01786; TonB-hemolactins; 1.  
 DR TIGRFAMs; TIGR01776; TonB-tbp-lbp; 1.

|                       |   |  |
|-----------------------|---|--|
| DR                    | InterPro; IPR010917; TonB recept C.                                   |  |
| DR                    | TIGRFAMS; TIGR01786; TonB-hemlactins; 1.                              |  |
| DR                    | PROSITE; PS01156; TONB_DEPENDENT_REC_2; UNKNOWN_1.                    |  |
| KW                    | Complete proteome; Hypothetical protein.                              |  |
| SQ                    | SEQUENCE 784 AA; 89543 MW; FAC10426E463FB2D CRC64;                    |  |
| Query Match           | 45.9%; Score 1927; DB 2; Length 784;                                  |  |
| Best Local Similarity | 48.1%; Pred. No. 3.4e-114;  |  |
| Matches               | 385; Conservative 135; Mismatches 202; Indels 78; Gaps 14;            |  |
| Qy                    | 28 ATETTPVKAIEKVRVKDQLNAPA-----TVERNLGRIOQEMIRNDKDLVRVSTDVGLS 82      |  |
| Db                    | 26 AQETT-----ELEQITVQESATAEVNKTSPVTSKATTIQEMVTRDLVRVTTDVGIS 81        |  |
| Qy                    | 83 DSGRHQKGFVAVGVEGRVGVSDVSLPDSSENSLYARYGNFNSLSLSDPELVNRIE 142        |  |
| Db                    | 82 DNGRFLKGFAMRGVEDNRVGVSDVSLPDSSENSLYARYGNFNSLSLSDPELIQIDI 141       |  |
| Qy                    | 143 AKGADSFNTSGALGGGVNTTQGHDLDDRFQVMMKNGYSSRRNREWTNTLFGVSN 202        |  |
| Db                    | 142 VRGSDSFNAGSGSLGGGVNTTLDPOHIVKAGNSVGALLRGSVASKNREWRTLIGYVG 201     |  |
| Qy                    | 203 DRVDAALLYSORGHETESAGERGYVPEGAGSGAIIKRGSRGIPDPSPKHKYHFLGKIAY 262   |  |
| Db                    | 202 EKFDALLMYSQRTGHEPKS-----RSGPPEFRYSSQHPDPVPTQRFHNYLAKQY 251        |  |
| Qy                    | 263 QINDKHRIGPSFNGOQHNTTIESYNLTASSREADDVNRNANLFYEWTPDSPNLSS 322       |  |
| Db                    | 252 QINDNQRIQUTNGTQGRVIDERSYTLMGQWREADDQERLNANDYYIYASTGWLAY 311       |  |
| Qy                    | 323 LKADFDYQTTKVAANVNGKS--FPTDYSTWTRNYNQDLENIYNRSMDFTRKFTLRMS 380     |  |
| Db                    | 312 SKFDLDYQKDLAAVYKGGHFTD-----AKELNEIYDRMKTVFTRGSVELNA 362           |  |
| Qy                    | 381 QPQLGQGHRLSLKTFASRREFENLNDYDFEERSVRTS-----SIQHPVKTINYG 433        |  |
| Db                    | 363 QPVHFGYHTLTIKGVYSQSDFKINQD-----RIGIGTNYDTQVHTTYQPIRTKQYG 416      |  |
| Qy                    | 434 FSLSDQIQWMDVFSSRADIRYDHTKMTPOELNAEC--HAC-DKTPPAANTYKWSGGFVGLA 491 |  |
| Db                    | 417 LSLKDHVRWMDTFSSHGLRHYDHTLKLKPELNAPCSKACLEBKPKPRTFTSTVFAGLE 476    |  |
| Qy                    | 492 AQLNQAHVGYDITSGYRVFNASEVYFTYNHSGNMLPNPNLKAERSTHTLSIQGRSEK 551     |  |
| Db                    | 477 AQLSPSWMLGVNISTGYRVFTASEMFFSTNAYGTWKSNSPLKPEKSINHTLSLKGNSK 536    |  |
| Qy                    | 552 GTLDANLYQNNRNFLSEBKLTTSBGVGTQOMNYYYGMCNSPYS--EKPWQWQNDKAR 610     |  |
| Db                    | 537 GLLDLTLYQNYRHLFPEQESL-----IQRTEMRY-----GRPYTQSQEQQWVNLDAK 586     |  |
| Qy                    | 611 IRGLELTGRNLVTKVASFVPEGWKLFGLGKLVAKSLGSDNLSLSTQPPKVIAGVDYESPS 670  |  |
| Db                    | 587 IYGVELTHVNLDQMIQAVIPQGFYALGYSKUSNNAASLISIQPLKIILGLDYBATN 646      |  |
| Qy                    | 671 EKQGVFSRLTYLGAKKAKDAQTYVYENKRGRT-----PLQKKVKD 710                 |  |
| Db                    | 647 GKWAIFNLRYLGEKRSADAK--VYEIKRRTREVTETDPTWQGITRCKRELYPDLST 704      |  |
| Qy                    | 711 YPMLNKSAFYDMGVFKYLAKNLTLRAGVNVNVRKTTWDSRLGLYSYSTTNAVDRDGK 770     |  |
| Db                    | 705 YKHLNKSFAFDFTFGYYKTTDDITFRAGIYNLFNKKYHTWDALRGINANSTLNSVDREGK 764  |  |
| Qy                    | 771 GLDVRASGRNYAVSLDWKF 790   |  |
| Db                    | 765 GLQRFYAPGRNYAASLEIRF 784  |  |
| RESULT 7              |   |  |
| ID                    | P72073  |  |
| AC                    | P72073; PRELIMINARY; PRT; 296 AA.                                     |  |
| DT                    | 01-FEB-1997 (TrEMBLrel. 02, Created)                                  |  |
| DT                    | 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)                     |  |

|                       |   |  |
|-----------------------|---|--|
| DT                    | 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)                   |  |
| DE                    | Hemoglobin receptor.  |  |
| GN                    | Name=hmbR;  |  |
| OS                    | Neisseria gonorrhoeae.  |  |
| OC                    | Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;           |  |
| OC                    | Neisseriaceae; Neisseria.   |  |
| OX                    | NCBI_TaxID=485;   |  |
| RN                    | [1]   |  |
| RP                    | SEQUENCE FROM N.A.  |  |
| RC                    | STRAIN=MS11;  |  |
| RA                    | Stojiljkovic I., Larson J., Hwa V., Anic S., So M.;                   |  |
| RL                    | Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.               |  |
| DR                    | EMBL; U40861; AAB18727.1; -   |  |
| DR                    | GO; GO:0019867; C:outer membrane; IEA.                                |  |
| DR                    | GO; GO:0004872; F:receptor activity; IEA.                             |  |
| DR                    | GO; GO:0005215; F:transporter activity; IEA.                          |  |
| DR                    | GO; GO:0006810; P:transport; IEA.                                     |  |
| KW                    | Receptor.   |  |
| SQ                    | SEQUENCE 296 AA; 32223 MW; 9520AE9B2A32947D CRC64;                    |  |
| Query Match           | 34.8%; Score 1463; DB 2; Length 296;                                  |  |
| Best Local Similarity | 39.8%; Pred. No. 3.2e-85;   |  |
| Matches               | 283; Conservative 0; Mismatches 1; Indels 0; Gaps 0;                  |  |
| Qy                    | 1 MKPLHMLPIAALVGSIFGNPVLAADEAAATETTPVKAIEKVRVKDQLNAPATVERVNLGR 60     |  |
| Db                    | 13 MKPLHMLPIAALVGSIFGNPVLAADEAAATETTPVKAIEKVRVKDQLNAPATVERVNLGR 72    |  |
| Qy                    | 61 IQEMIRDNKDLVRYSTDVGLSDSGRHQKGFVAVGVEGRVGVSDVSLPDSSENSLYA 120       |  |
| Db                    | 73 IQEMIRDNKDLVRYSTDVGLSDSGRHQKGFVAVGVEGRVGVSDVSLPDSSENSLYA 132       |  |
| Qy                    | 121 RYGNFNSSRLSIDPELVNRIEIAKGADSFNTSGALGGGVNTTQGHDLDDRFQVSN 180       |  |
| Db                    | 133 RYGNFNSSRLSIDPELVNRIEIAKGADSFNTSGALGGGVNTTQGHDLDDRFQVSN 192       |  |
| Qy                    | 181 MKNGYSSRRNREWTNTLFGVSNDRVDAALLYSORGHETESAGERGYVPEGAGSGAIIIRG 240  |  |
| Db                    | 193 MKNGYSSRRNREWTNTLFGVSNDRVDAALLYSORGHETESAGERGYVPEGAGSGANIRG 252   |  |
| Qy                    | 241 SSRGIDPSPKHKYHFLGKIAYQINDKHRIGPSFNGOQHNTT 284                     |  |
| Db                    | 253 SSRGIDPSPKHKYHFLGKIAYQINDKHRIGPSFNGOQHNTT 296                     |  |
| RESULT 8              |   |  |
| ID                    | Q9RGD8  |  |
| AC                    | Q9RGD8; PRELIMINARY; PRT; 209 AA.                                     |  |
| DT                    | 01-MAY-2000 (TrEMBLrel. 13, Created)                                  |  |
| DT                    | 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)                     |  |
| DT                    | 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)                   |  |
| DE                    | HmbR (Fragment).  |  |
| GN                    | Name=hmbR;  |  |
| OS                    | Neisseria meningitidis.   |  |
| OC                    | Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;           |  |
| OC                    | Neisseriaceae; Neisseria.   |  |
| OX                    | NCBI_TaxID=487;   |  |
| RN                    | [1]   |  |
| RP                    | SEQUENCE FROM N.A.  |  |
| RC                    | STRAIN=IR2855;  |  |
| RX                    | MEDLINE=20096688; PubMed=10629191; DOI=10.1128/JB.182.2.439-447.2000; |  |
| RA                    | Zhu W., Hunt D.J., Richardson A.R., Stojiljkovic I.;                  |  |
| RT                    | "Use of Heme Compounds as Iron Sources by Pathogenic Neisseriae       |  |
| RT                    | requires the product of the hemo gene.";                              |  |
| RL                    | J. Bacteriol. 182:439-447(2000).                                      |  |
| DR                    | EMBL; AFI33695; AAF24746.1; -   |  |
| DR                    | GO; GO:0016020; C:membrane; IEA.                                      |  |
| DR                    | GO; GO:0004872; F:receptor activity; IEA.                             |  |
| DR                    | GO; GO:0005215; F:transporter activity; IEA.                          |  |
| DR                    | GO; GO:0006810; P:transport; IEA.                                     |  |
| FT                    | NON_TER 209 209   |  |
| SQ                    | SEQUENCE 209 AA; 22683 MW; DE5CD5FF0AD40759 CRC64;                    |  |

```
Db 361 DBIYNRMDTRFKRITRLDRLSHPLQLGGRRHLSFKTFASRRDRFENLNRDDYCFSGRVVR 420
Qy 419 TTSSIOHPVKTNTNYGSLSDQIQWNVDFSSRADIRYDHTKMTPOELNAEACHADCTPPAA 478
Db 421 TTSSIOHPVKTNTNYGSLSDQIQWNVDFSSRADIRYDHTKMTPOELNAEACHADCTPPAA 480
Qy 479 NTYKMGSGFVGLAAQLNQAHWGYDITSGYRVPNASEVFTYNHGSGNWLPPNPKAERS 538
Db 481 NTYKMGSGFVGLAAQLNQAHWGYDITSGYRVPNASEVFTYNHGSGNWLPPNPKAERS 540
Qy 539 THTLSLQGRSEKGTLDANLYQNNYRNFLESEOKLTSGDVGCTQNNYYGCMSPYSEK 598
Db 541 THTLSLQGRSEKGTLDANLYQNNYRNFLESEOKLTSGDVGCTQNNYYGCMSPYSEK 600
Qy 599 PEMQONIDKARIRGLTGRNLNVTKVASFPVPEGWKLFGSLGVAKSKLSDGNSLLSTOPP 658
Db 601 PEMQONIDKARIRGLTGRNLNVTKVASFPVPEGWKLFGSLGVAKSKLSDGNSLLSTOPP 660
Qy 659 KVIAGDYESPSEKMGVFSRLTYLGAKKAKDAQYTYVYENKGRGTPLOKKVKDYPWLKNSA 718
Db 661 KVIAGDYESPSEKMGVFSRLTYLGAKKAKDAQYTYVYENKGRGTPLOKKVKDYPWLKNSA 720
Qy 719 YVDMYGFYKLAQNLTLAGVNVFNKRYTTWDSLRGLYSYSTTNAVDRDGLDRYRAS 778
Db 721 YVDMYGFYKLAQNLTLAGVNVFNKRYTTWDSLRGLYSYSTTNAVDRDGLDRYRAS 780
Qy 779 GRNAYVSLDWKF 790
Db 781 GRNAYVSLDWKF 792

RESULT 5
ID P72085 PRELIMINARY; PRT; 791 AA.
AC P72085;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hemoglobin receptor.
GN Name=hmbR;
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1] _
RP SEQUENCE FROM N.A.
RC STRAIN=Serogroup B;
RA Stojiljkovic I., Larson J., Hwa V., Anic S., So M.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40860; AAB18729.1; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR010949; TonB hemolactins.
DR TIGRFAMs; TIGR01786; TonB-hemolactins; 1.
KW Receptor.
SQ SEQUENCE 791 AA; 89281 MW; 12B957A15232B39C CRC64;

Query Match 87.3%; Score 3665.5; DB 2; Length 791;
Best Local Similarity 86.9%; Pred. No. 5.2e-225;
Matches 687; Conservative 44; Mismatches 59; Indels 1; Gaps 1;

Qy 1 MKPLHMLPAAVLGSGTFCGNPLVAADAEATETTPVKAEIKEVKVKDQNLNAPATVERVNLGR 60
Db 1 MKPLHMLPAAVLGSGTFCGNPLVAADAEATETTPVKAEIKEVKVKDQNLNAPATVERVNLGR 60
Qy 61 IQQEMIRNDKLVRYSTDVGLSDGRGHQKGFVRGVEGNVGVSDIGVSLPDSSENSLYA 120
Db 61 IQQEMIRNDKLVRYSTDVGLSDGRGHQKGFVRGVEGNVGVSDIGVSLPDSSENSLYA 120
Qy 121 RYGNFNSSRLSIDPELVNRNIETAKGADSPNTGSGALGGGVNTYQTLQGRDLLLPERQFGVM 180
Db 121 RYGNFNSSRLSIDPELVNRNIETAKGADSPNTGSGALGGGVNTYQTLQGRDLLLPERQFGVM 180
```

```
RESULT 6
Q9CMS1
ID Q9CMS1 PRELIMINARY; PRT; 784 AA.
AC Q9CMS1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein PM0741.
GN OrderedLocusNames=PM0741;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RA MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006111; AAK02825.1; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR010949; TonB hemolactins.
```

```

OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8013 clone 6;
RX MEDLINE=95302964; PubMed=7783623;
RA Stojiljkovic I., Hwa V., de Saint Martin L., O'Gaora P., Nassif X.,
RA Heffron F., So M.;
RT "The Neisseria meningitidis haemoglobin receptor: its role in iron
RT utilization and virulence.";
RL Mol. Microbiol. 15:531-541(1995).
DR EMBL; U18558; AA80205.1; -.
DR PIR; S61335; S61335.
DR GO; GO:0019867; C:outer membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR010949; TonB hemolactrns.
DR TIGRFAMs; TIGR01786; TonB-hemolactrns; 1.
KW Receptor.
SQ SEQUENCE 792 AA; 89405 MW; 52381200297B37A2 CRC64;

Query Match 90.8%; Score 3812; DB 2; Length 792;
Best Local Similarity 91.2%; Pred. No. 2.4e-234;
Matches 722; Conservative 25; Mismatches 43; Indels 2; Gaps 2;

Qy 1 MKPLHMLPIAALVGSIFGNPVLAADEAATETTPVKAEIKEVVRKQDLNAPATVERVNLGR 60
Db 1 MKPLQMLPVAALVGSIFGNPVADEAATETTPVKAQVAVKVGQRNAPAAVERVNLNR 60

Qy 61 IQQMIRDNDKLVRYSTDVGLSDSGRHQKGFVAVRGEVGNRVGYSIDGVSLPDSSENSLYA 120
Db 61 IQQMIRDNDKLVRYSTDVGLSDSGRHQKGFVAVRGEVGNRVGYSIDGVSLPDSSENSLYA 120

Qy 121 RYGNFNSRSLSIDPELVNRNIDIVKGADSFNTGSGALGGGVNYOTLQGHDLILLDDRQFGVM 180
Db 121 RYGNFNSRSLSIDPELVNRNIDIVKGADSFNTGSGALGGGVNYOTLQGHDLILLDDRQFGVM 180

Qy 181 MKNGYSRNRNREWTNTLFGVSNDRVDAALYSQRRGHETESAGRGYPVEGAGSGAIRG 240
Db 181 MKNGYSRNRNREWTNTLFGVSNDRVDAALYSQRRGHETESAGRGYPVEGAGSGAIRG 240

Qy 241 SSRGIPDPSPKHKNFLGKIAYQINDKHRIKPSFNGOQHNYTIEESYNTLASSWRADD 300
Db 241 SARGIPDPSPKHKNFLGKIAYQINDKHRIKPSFNGOQHNYTIEESYNTLASSWRADD 300

Qy 301 VNRNRNANLFYEWTDPDSNWLSSLKADFDYOTTKVAAVNNKGSFP-TDYSTWTRNYNOKDL 359
Db 301 VNRNRNANLFYEWTDPDSNWLSSLKADFDYOTTKVAAVNNKGSFP-TDYSTWTRNYNOKDL 359

Qy 360 ENIYNRSMDFRFRFTLRMDSPQLQ-LGGQHRISLKTFFASRRFENLNRRDDYFSEVSR 418
Db 360 ENIYNRSMDFRFRFTLRMDSPQLQ-LGGQHRISLKTFFASRRFENLNRRDDYFSEVSR 418

Qy 419 TTSSIQHPVKNTNYGFSLSQIQNDVFSRADIRYDHTKMTPOELNAEACDKTTPAA 478
Db 419 TTSSIQHPVKNTNYGFSLSQIQNDVFSRADIRYDHTKMTPOELNAEACDKTTPAA 478

Qy 479 NTYKWSGFGVLAQLNQAHVGVYDITSGYRVPNASEVYFTYNHGSGNWLPNPLKAERS 538
Db 479 NTYKWSGFGVLAQLNQAHVGVYDITSGYRVPNASEVYFTYNHGSGNWLPNPLKAERS 538

Qy 539 TTHTLSLQGRSEKGTLDANLYQNNYRNFLSEEQKLTTSGDVGCTQMYYYGMCSPYSEK 598
Db 539 TTHTLSLQGRSEKGTLDANLYQNNYRNFLSEEQKLTTSGDVGCTQMYYYGMCSPYSEK 598

Qy 599 PEWQMONIDKARIGELTGLNLYTKVASVPEGKMLFGSLGYAKSLGSDNSLLSTOPP 658
Db 599 PEWQMONIDKARIGELTGLNLYTKVASVPEGKMLFGSLGYAKSLGSDNSLLSTOPP 658

Qy 601 LEWQMONIDKARIGELTGLNLYTKVASVPEGKMLFGSLGYAKSLGSDNSLLSTOPL 660
Db 601 LEWQMONIDKARIGELTGLNLYTKVASVPEGKMLFGSLGYAKSLGSDNSLLSTOPL 660

Qy 659 KVIAGDYESPSEKGVFSRLTYLGAKKAKADAQYTVYENKGRGTPLOKKVKDYPWLKNSA 718
Db 659 KVIAGDYESPSEKGVFSRLTYLGAKKAKADAQYTVYENKGRGTPLOKKVKDYPWLKNSA 718

```

```

Db 121 RYGNFSSRLSDPELVRIEIVKGDSTSGALGGVNYQTQGRDLLDDRFQGYM 180
Qy 181 MKNGYSSRNREWTNTLFGVSNDRVDAAALYSORRGHETESAGRGYPVEGAGSIAIRG 240
Db 181 MKNGYSTNRREWTNTLFGVSNDRVDAAALYSORRGHETESAGRGYAVEGSGANIRG 240
Qy 241 SSGIPDPKHKYHNFGLGIAQINDKHRIGPSFNGQQGHNTTIBSYNLTASSWREADD 300
Db 241 SARGIPDSSKHKYHSLFLGIAQINDKHRIASLNGQQGHNTTIBSYNLTASSWREADD 300
Qy 301 VNRRNANLFYEWTPDSSNLSLKADFDYQTKVAAVNNKGSFPDPTDYSYTWNTYNNQKOLE 360
Db 301 VNRRNANLFYEWTPDSSNLSLKADFDYQTKVAAVNNKGSFPDPTDYSYTWNTYNNQKOLD 360
Qy 361 NIYNRSMDFRKFRTLRMDSDPLQL-GGOHRLSLKTFASRRFENLNDDYFYSERVST 419
Db 361 EYINRSMDFRKFRTLRDLSHPQLGGGRHRLSKFTFVSRRDFENLNDDYFYSGRVVRT 420
Qy 420 TSSIOHPVKTNTYGFSLSDQIQWNDVFSRRADIRYDHTKMTPOELNAEACHADCKTPPAAN 479
Db 421 TSSIOHPVKTNTYGFSLSDQIQWNDVFSRRADIRYDHTKMTPOELNAEACHADCKTPPAAN 480
Qy 480 TYKGSFGVGLAAQLNQAHVGYDITSGYRVPNASEVYFTYNHSGSNLPLPNLKAERST 539
Db 481 TYKGSFGVGLAAQLNQAHVGYDITSGYRVPNASEVYFTYNHSGSNLPLPNLKAERST 540
Qy 540 THTLSQGRSEKGTLDANLYQNNYRNFLESEOKLTTSGDVGCTOMNYNYGMSNPYSEKP 599
Db 541 THTLSQGRSEKGTLDANLYQNNYRNFLESEOKLTTSGPGCTEENAYYICSDPYKEKL 600
Qy 600 EQWQONIDKARIRGLTGRNLNVTKVASFVPEGWKLFGLGAKKADQYTYVENKGRGTPLOKKYDYPWLKNSAY 719
Db 601 DMQKNIDKARIRGLTGRNLNVTKVASFVPEGWKLFGLGAKKADQYTYVENKGRGTPLOKKYDYPWLKNSAY 720
Qy 720 VFDYMGFYKLAQNLTLRAGVYNVFNKRYTTWDSLRGLYSYSTTNAVDRDCKGLDRYRAGS 779
Db 721 VFDYMGFYKPAKNLTLRAGVYNLFRNKRYTTWDSLRGLYSYSTTNAVDRDCKGLDRYRAGP 780
Qy 780 RNYAVSLDWKF 790
Db 781 RNYAVSLEWKF 791

```

## RESULT 2

```

P72084
ID P72084 PRELIMINARY; PRT; 791 AA.
AC P72084;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hemoglobin receptor.
GN Name=hmbR;
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serogroup A;
RA Stojiljkovic I., Larson J., Hwa V., Anic S., So M.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U40859; AAB18728.1; -.
DR GO; GO:0019867; C:outer membrane; IEA.
DR GO; GO:004872; F:receptor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transpor; IEA.
DR InterPro; IPR010949; TonB hemolactrns.
DR TIGRFAMs; TIGR01786; TonB-hemolactrns; 1.

```

```

KW Receptor.
SQ SEQUENCE 791 AA; 89359 MW; A82C4721DCAAGED0 CRC64;

Query Match 90.9%; Score 3818.5; DB 2; Length 791;
Best Local Similarity 90.9%; Pred. No. 9.1e-235;
Matches 719; Conservative 25; Mismatches 46; Indels 1; Gaps 1;

Qy 1 MKPLHMLPIAALVGSIFGNPVLAADEAAATETTPVKAETKEVRVKDQLNAPATVERVNLGR 60
Db 1 MKPLQMLPIAALVGSIFGNPVLAADEAAATETTPVKAETKAVRVKQORAPAAVERVNLNR 60
Qy 61 IQQEMIRNDKOLVRYSTDVGLSDSRGHQKGFVAVRGVEGNRVGVSIDGVSPLDSEENSLYA 120
Db 61 IQQEMIRNDKOLVRYSTDVGLSDSRGHQKGFVAVRGVEGNRVGVSIDGVNLPDSEENSLYA 120
Qy 121 RYGNFSSRLSDPELVRIEIVKGDSTSGALGGVNYQTQGRDLLDDRFQGYM 180
Db 121 RYGNFSSRLSDPELVRIEIVKGDSTSGALGGVNYQTQGRDLLDDRFQGYM 180
Qy 181 MKNGYSSRNREWTNTLFGVSNDRVDAAALYSORRGHETESAGRGYPVEGAGSIAIRG 240
Db 181 MKNGYSTNRREWTNTLFGVSNDRVDAAALYSORRGHETESAGRGYAVEGAGKETNIRG 240
Qy 241 SSGIPDPKHKYHNFGLGIAQINDKHRIGPSFNGQQGHNTTIBSYNLTASSWREADD 300
Db 241 SARGIPDSSKHKYHNFGLGIAQINDKHRIASLNGQQGHNTTIBSYNLTASSWREADD 300
Qy 301 VNRRNANLFYEWTPDSSNLSLKADFDYQTKVAAVNNKGSFPDPTDYSYTWNTYNNQKOLE 360
Db 301 VNRRNANLFYEWTPDSSNLSLKADFDYQTKVAAVNNKGSFPDPTDYSYTWNTYNNQKOLE 360
Qy 361 NIYNRSMDFRKFRTLRMDSDPLQL-GGOHRLSLKTFASRRFENLNDDYFYSERVST 419
Db 361 EYINRSMDFRKFRTLRDLSHPQLGGGRHRLSKFTFVSRRDFENLNDDYFYSGRVVRT 420
Qy 420 TSSIOHPVKTNTYGFSLSDQIQWNDVFSRRADIRYDHTKMTPOELNAEACHADCKTPPAAN 479
Db 421 TSSIOHPVKTNTYGFSLSDQIQWNDVFSRRADIRYDHTKMTPOELNAEACHADCKTPPAAN 480
Qy 480 TYKGSFGVGLAAQLNQAHVGYDITSGYRVPNASEVYFTYNHSGSNLPLPNLKAERST 539
Db 481 TYKGSFGVGLAAQLNQAHVGYDITSGYRVPNASEVYFTYNHSGSNLPLPNLKAERST 540
Qy 540 THTLSQGRSEKGTLDANLYQNNYRNFLESEOKLTTSGDVGCTOMNYNYGMSNPYSEKP 599
Db 541 THTLSQGRSEKGTLDANLYQNNYRNFLESEOKLTTSGPGCTEENAYYICSDPYKEKL 600
Qy 600 EQWQONIDKARIRGLTGRNLNVTKVASFVPEGWKLFGLGAKKADQYTYVENKGRGTPLOKKYDYPWLKNSAY 719
Db 601 DMQKNIDKARIRGLTGRNLNVTKVASFVPEGWKLFGLGAKKADQYTYVENKGRGTPLOKKYDYPWLKNSAY 720
Qy 720 VFDYMGFYKLAQNLTLRAGVYNVFNKRYTTWDSLRGLYSYSTTNAVDRDCKGLDRYRAGS 779
Db 721 VFDYMGFYKPAKNLTLRAGVYNLFRNKRYTTWDSLRGLYSYSTTNAVDRDCKGLDRYRAGP 780
Qy 780 RNYAVSLDWKF 790
Db 781 RNYAVSLEWKF 791

RESULT 3
ID Q51104 PRELIMINARY; PRT; 792 AA.
AC Q51104;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hemoglobin receptor.
GN Name=hmbR;

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2005, 14:49:54 ; Search time 178 seconds  
(without alignments)  
2272.711 Million cell updates/sec

Title: US-09-665-358-8  
Perfect score: 4200  
Sequence: 1 MKPLHMLPIAALVGSIFGNP.....GLDRYASGRNYAVSLDWKF 790

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | ID           | Description            |
|------------|--------|-------------|--------|--------------|------------------------|
| 1          | 3867.5 | 92.1        | 791    | 2 Q9JYA8     | Q9JYA8 neisseria m     |
| 2          | 3818.5 | 90.9        | 791    | 2 P72084     | P72084 neisseria m     |
| 3          | 3812   | 90.8        | 792    | 2 Q51104     | Q51104 neisseria m     |
| 4          | 3800   | 90.5        | 792    | 2 Q9RG39     | Q9RG39 neisseria m     |
| 5          | 3665.5 | 87.3        | 791    | 2 P72085     | P72085 neisseria m     |
| 6          | 1927   | 45.9        | 784    | 2 Q9CMS1     | Q9CMS1 pasteurella     |
| 7          | 1463   | 34.8        | 296    | 2 P72073     | P72073 neisseria g     |
| 8          | 1003   | 23.9        | 209    | 2 Q9RQD8     | Q9RQD8 neisseria m     |
| 9          | 857    | 20.4        | 763    | 2 Q9GB30     | Q9GB30 haemophilus     |
| 10         | 854    | 20.3        | 759    | 2 Q934H9     | Q934H9 haemophilus     |
| 11         | 850.5  | 20.2        | 760    | 2 Q934H8     | Q934H8 haemophilus     |
| 12         | 833    | 19.8        | 777    | 2 Q9ALL8     | Q9ALL8 pasteurella     |
| 13         | 827    | 19.7        | 777    | 2 Q93295     | Q93295 pasteurella     |
| 14         | 792.5  | 18.9        | 972    | 1 HGBA HAEU  | HGBA HAEU haemophilus  |
| 15         | 786.5  | 18.7        | 810    | 1 H9UB NEIWA | H9UB NEIWA neisseria m |
| 16         | 786.5  | 18.7        | 810    | 1 H9UB NEIMC | H9UB NEIMC neisseria m |
| 17         | 780.5  | 18.6        | 946    | 2 Q6JPV5     | Q6JPV5 actinobacil     |
| 18         | 780.5  | 18.6        | 946    | 2 Q6JPV7     | Q6JPV7 actinobacil     |
| 19         | 775.5  | 18.5        | 993    | 1 H9CB HAELN | H9CB HAELN haemophilus |
| 20         | 772.5  | 18.4        | 948    | 2 Q941A7     | Q941A7 actinobacil     |
| 21         | 767.5  | 18.3        | 1063   | 1 H9PB HAEIN | H9PB HAEIN haemophilus |
| 22         | 767.5  | 18.3        | 1067   | 1 H9BB HAELN | H9BB HAELN haemophilus |
| 23         | 766    | 18.2        | 1066   | 1 H9PC HAELN | H9PC HAELN haemophilus |
| 24         | 759    | 18.1        | 1084   | 1 H9P3 HAELN | H9P3 HAELN haemophilus |
| 25         | 755.5  | 18.0        | 1013   | 1 H9BA HAELN | H9BA HAELN haemophilus |
| 26         | 753.5  | 17.9        | 967    | 2 Q9L6A7     | Q9L6A7 pasteurella     |
| 27         | 752.5  | 17.9        | 999    | 1 H9P2 HAELN | H9P2 HAELN haemophilus |
| 28         | 749    | 17.8        | 999    | 1 H9PB HAELN | H9PB HAELN haemophilus |
| 29         | 745.5  | 17.8        | 1023   | 2 Q932Z6     | Q932Z6 actinobacil     |
| 30         | 745.5  | 17.8        | 1023   | 2 Q93TF7     | Q93TF7 actinobacil     |
| 31         | 743.5  | 17.7        | 1023   | 2 Q934C2     | Q934C2 actinobacil     |

|    |       |      |      |              |                        |
|----|-------|------|------|--------------|------------------------|
| 32 | 736   | 17.5 | 999  | 1 HGP4 HAEIN | O57408 haemophilus     |
| 33 | 734.5 | 17.5 | 1023 | 2 Q93TF8     | Q93TF8 actinobacil     |
| 34 | 730.5 | 17.4 | 971  | 2 Q71RX0     | Q71RX0 pasteurella     |
| 35 | 725.5 | 17.3 | 963  | 2 Q9CNX2     | Q9CNX2 pasteurella     |
| 36 | 725   | 17.3 | 989  | 2 Q9CNT9     | Q9CNT9 pasteurella     |
| 37 | 712.5 | 17.0 | 764  | 2 O68881     | O68881 pseudomonas     |
| 38 | 707   | 16.8 | 1077 | 1 HGP4 HAEIN | HGP4 HAEIN haemophilus |
| 39 | 706.5 | 16.8 | 764  | 2 Q9HV88     | Q9HV88 pseudomonas     |
| 40 | 703.5 | 16.8 | 857  | 2 Q887K8     | Q887K8 pasteurella     |
| 41 | 695   | 16.5 | 997  | 2 Q9CNT8     | Q9CNT8 haemophilus     |
| 42 | 695   | 16.5 | 1046 | 1 HHUA HAEIN | HHUA HAEIN haemophilus |
| 43 | 685.5 | 16.3 | 944  | 1 LBPA NEIMA | LBPA NEIMA neisseria m |
| 44 | 669.5 | 15.9 | 943  | 2 Q50952     | Q50952 neisseria g     |
| 45 | 666.5 | 15.9 | 715  | 2 Q6LRY0     | Q6LRY0 photobacter     |

ALIGNMENTS

RESULT 1

Q9JYA8  
ID Q9JYA8 PRELIMINARY; PRT; 791 AA.  
AC Q9JYA8; 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Hemoglobin receptor.  
GN OrderedLocustNames=NM1668;  
OS Neisseria meningitidis (serogroup B).  
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
OC Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=491;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MC58 / Serogroup B;  
RX MEDLINE=2017575; PubMed=10710307; DOI=10.1126/science.287.5459.1809;  
RA Tettelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,  
RA Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,  
RA Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R.T., Peterson J.D.,  
RA Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,  
RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,  
RA Clifton H., Clark E.B., Cotton M.D., Uitterback T.R., Khouri H.M.,  
RA Qin H., Vamathevan J.J., Gill J., Scariati V., Masignani V., Fizza M.,  
RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,  
RA Venter J.C.;  
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.";  
RL Science 287:1809-1815(2000).  
DR EMBL; AR002517; AAF42017.1; -.  
DR PIR; F81056; F81056.  
DR TIGR; NM1668; -.  
DR GO; GO:0019867; C:outer membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR010949; TonB hemolactins.  
DR TIGRFAMs; TIGR01786; TonB-hemolactins; 1.  
KW Complete proteome; Receptor.  
SQ SEQUENCE 791 AA; 89218 MW; F055F59A6A36ED3A CRC64;

Query Match 92.1%; Score 3867.5; DB 2; Length 791;  
Best Local Similarity 92.3%; Pred. No. 6.9e-238;  
Matches 730; Conservative 20; Mismatches 40; Indels 1; Gaps 1;

|    |     |   |     |
|----|-----|---|-----|
| Qy | 1   | MKPLHMLPIAALVGSIFGNPVLAADEAATTPVKAIKEVRVKDQLNAPATVERVNLGR   | 60  |
| Db | 1   | MKPLQMLPIAALVGSIFGNPVLAADEAATTPVKAIKAVRVKQGNAPAAVERVNLNR    | 60  |
| Qy | 61  | IQCEMIRNDKDLVRYSTDVGLSDSGRQKGFVAVRGVGNRVGVSIDGVLSPDSENSIYA  | 120 |
| Db | 61  | IKQCEMIRNDKDLVRYSTDVGLSDSGRQKGFVAVRGVGNRVGVSIDGVLSPDSENSIYA | 120 |
| Qy | 121 | RYGNFNSRLSIDPELVNRNIEIAKGAUSFNTGSGALOGGVNYQTLOGHDLDDRRQGVN  | 180 |

Db 361 EYNRSMDTRFKFTLRDLSHPLQLGGGRHRLSFKTFVSRDFENLRDDYFSGRVRT 420  
Qy 420 TSSIQHPVKTNYGFSLSDOIQWNVFSSRADIRYDHTKMTPOELNAEACHADCKTPPAAN 479  
Db 421 TSSIQHPVKTNYGFSLSDOIQWNVFSSRADIRYDHTKMTPOELNAEACHADCKTPPAAN 480  
Qy 480 TYKWSGFGVGLAAQLNAQWVGVDITSGYRVPNASEVFTYNHGSGNWLNPENLKAERST 539  
Db 481 TYKWSGFGVGLAAQLNAQWVGVDITSGYRVPNASEVFTYNHGSGNWLNPENLKAERST 540  
Qy 540 THTLSQGRSEKGTLDANLYQNNYRNFLESEOKLTTSQDVGCTQNNYYGMCSPYSEKP 599  
Db 541 THTLSQGRSEKGTLDANLYQNNYRNFLESEOKLTTSQDVGCTQNNYYGMCSPYSEKP 600  
Qy 600 EYQWQNDIKARIRGLTGRNLNVTYKASFPVPEGWKLFGLSGYAKSKLSDNSLLSTQPK 659  
Db 601 EYQWQNDIKARIRGLTGRNLNVTYKASFPVPEGWKLFGLSGYAKSKLSDNSLLSTQPK 660  
Qy 660 VIAGDYSPSEKMGVFSRLTYLGAKKADQVTVYENKGRGTPLQKKVKDYPWLNKSA 719  
Db 661 VIAGDYSPSEKMGVFSRLTYLGAKKADQVTVYENKGRGTPLQKKVKDYPWLNKSA 720  
Qy 720 VFDYMGFYKAKNLTLAGVYVNFNRKYTTWDSLRGLYSYSTTNAVDRDCKGLDRYRAG 779  
Db 721 VFDYMGFYKAKNLTLAGVYVNFNRKYTTWDSLRGLYSYSTTNAVDRDCKGLDRYRAG 780  
Qy 780 RNYAVSLDWKF 790  
Db 781 RNYAVSLEWKF 791

## RESULT 2

hemoglobin receptor precursor - Neisseria meningitidis  
C:Species: Neisseria meningitidis  
C:Date: 27-Apr-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S61335  
R:Stojiljkovic, I.; Hwa, V.; de Saint Martin, L.; O'Gaora, P.; Nassif, X.; Heffron, F.; Mol. Microbiol. 15, 531-541, 1995  
A:Title: The Neisseria meningitidis haemoglobin receptor: its role in iron utilization  
A:Reference number: S61335; PMID:95302964; PMID:7783623  
A:Accession: S61335  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-792 <STO>  
A:Cross-references: UNIPROT:Q51104; EMBL:U18558; NID:U687639; PIDN:AAA80205.1; PID:96876  
C:Superfamily: ferrichrome-iron receptor; tonB-dependent receptor amino-terminal homolog  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-792/Product: hemoglobin receptor #status predicted <MAT>  
F:68-206/Domain: tonB-dependent receptor amino-terminal homology <TNN>

Query Match 90.8%; Score 3812; DB 2; Length 792;  
Best Local Similarity 91.2%; Pred. No. 1.3e-242;  
Matches 722; Conservative 25; Mismatches 43; Indels 2; Gaps 2;

Qy 1 MKPLHMLPIAALVGSIFGPNVLAADSAATETTPVKAEIKEVVRKQDLNAPATVERVNLGR 60  
Db 1 MKPLQMLPIAALVGSIFGPNVLAADSAATETTPVKAEIKEVVRKQDLNAPATVERVNLGR 60  
Qy 61 IQQEMIRNDKOLRYSTDVGLSDSGRHKQGFVAVRGVEGNRVGSDIGVSLPDSSENSLYA 120  
Db 61 IQQEMIRNDKOLRYSTDVGLSDSGRHKQGFVAVRGVEGNRVGSDIGVSLPDSSENSLYA 120  
Qy 121 RYGNFNSSRLSDPELVRNIETAKGADSFNTGSGALGGGVNTQTQGHDLDDRRQFGVM 180  
Db 121 RYGNFNSSRLSDPELVRNIDIVKGAADFNTGSGALGGGVNTQTQGHDLDDRRQFGVM 180  
Qy 181 MKNYSSRNRENTNTLFGVSDNRVDAAALYSORRHETESAGERYPVEGAGSGAIIIRG 240  
Db 181 MKNYSSRNRENTNTLFGVSDNRVDAAALYSORRHETESAGERYPVEGAGSGAIIIRG 240  
Qy 241 SSRGIPDPSPKHYHNFGLGIAYQINDKHRIQSGFNGQQGHNTYIEESYNLTASWREADD 300

Db 241 SARGIPDPSPKHYHNFGLGIAYQINDKHRIQSGFNGQQGHNTYIEESYNLTASWREADD 300  
Qy 301 VNRNRNANLFYEWTPDPSNWLSSLKADPDYQTTKVAANVKNKSEF-TDYSTVTRVYNQDL 359  
Db 301 VNRNRNANLFYEWTPDPSNWLSSLKADPDYQTTKVAANVKNKSEF-TDYSTVTRVYNQDL 360  
Qy 360 ENIYNRSMDTRFKFTLRDLSHPLQLGGGRHRLSFKTFVSRDFENLRDDYFSGRVVR 418  
Db 361 ENIYNRSMDTRFKFTLRDLSHPLQLGGGRHRLSFKTFVSRDFENLRDDYFSGRVVR 420  
Qy 419 TSSIQHPVKTNYGFSLSDOIQWNVFSSRADIRYDHTKMTPOELNAEACHADCKTPPA 478  
Db 421 TSSIQHPVKTNYGFSLSDOIQWNVFSSRADIRYDHTKMTPOELNAEACHADCKTPPA 480  
Qy 479 NTYKWSGFGVGLAAQLNAQWVGVDITSGYRVPNASEVFTYNHGSGNWLNPENLKAERS 538  
Db 481 NTYKWSGFGVGLAAQLNAQWVGVDITSGYRVPNASEVFTYNHGSGNWLNPENLKAERS 540  
Qy 539 THTLSQGRSEKGTLDANLYQNNYRNFLESEOKLTTSQDVGCTQNNYYGMCSPYSEK 598  
Db 541 THTLSQGRSEKGTLDANLYQNNYRNFLESEOKLTTSQDVGCTQNNYYGMCSPYSEK 600  
Qy 599 PEWQWQNDIKARIRGLTGRNLNVTYKASFPVPEGWKLFGLSGYAKSKLSDNSLLSTOPP 658  
Db 601 LEWQWQNDIKARIRGLTGRNLNVTYKASFPVPEGWKLFGLSGYAKSKLSDNSLLSTOPP 660  
Qy 659 KVIAGDYSPSEKMGVFSRLTYLGAKKADQVTVYENKGRGTPLQKKVKDYPWLNKSA 718  
Db 661 KVIAGDYSPSEKMGVFSRLTYLGAKKADQVTVYENKGRGTPLQKKVKDYPWLNKSA 720  
Qy 719 YVDMYGFYKAKNLTLAGVYVNFNRKYTTWDSLRGLYSYSTTNAVDRDCKGLDRYRAS 778  
Db 721 YVDMYGFYKAKNLTLAGVYVNFNRKYTTWDSLRGLYSYSTTNAVDRDCKGLDRYRAS 780  
Qy 779 GRNVAVSLDWKF 790  
Db 781 SRNVAVSLDWKF 792

RESULT 3  
A81965  
hemoglobin-haptoglobin-utilization protein NMA0474 [imported] - Neisseria meningitidis  
C:Species: Neisseria meningitidis  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004  
C:Accession: A81965  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, J.; Holroyd, S.; Jagsis, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A:Reference number: A81775; MUID:20222556; PMID:10761919  
A:Accession: A81965  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-810 <PAR>  
A:Cross-references: UNIPROT:Q9JWA2; GB:ALI62753; GB:ALI57959; NID:g7379120; PIDN:CAB83763  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: hpuB; NMA0474

Query Match 18.7%; Score 786.5; DB 2; Length 810;  
Best Local Similarity 26.9%; Pred. No. 1.2e-43;  
Matches 232; Conservative 155; Mismatches 349; Indels 125; Gaps 29;

Qy 3 PLHMLPIAALVGSIFGPNVLAADSAATETTPVKAE-IKEVRV----KQOLNAPATVERVN 57  
Db 2 PIPFVPLAALVGSIFGPNVLAADSAATETTPVKAE-----PQSAQTLINEITVTGHTKQGLGEEKIRRT 56  
Qy 58 LGRIOEMIRNDKOLRYSTDVGLSDSGR-HQKGFVAVRGVEGNRVGSDIGVSLPDSSEEN 116  
Db 57 LDKL---LVNDEHDLVRYDPGISVVEGGRAGSNGTIRGVKDRVAINVGLQAQAESRS 113  
Qy 117 ---SLYARYGNFNSSRLSDPELVRNIETAKGADSFNTGSGALGGGVNTQTQGHDL 172



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2005, 14:49:54 ; Search time 43 Seconds  
(without alignments)  
1767.703 Million cell updates/sec

Title: US-09-665-358-8  
Perfect score: 4200  
Sequence: 1 MKPLHMLPIAALVGSIFGNP.....GLDRYASGRNYAVSLDKWF 790

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:.\*  
1: Pir1.\*  
2: Pir2.\*  
3: Pir3.\*  
4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | ID       | Description        |
|------------|--------|-------------|--------|----------|--------------------|
| 1          | 3857.5 | 92.1        | 791    | 2 F81056 | hemoglobin recepto |
| 2          | 3812   | 90.8        | 792    | 2 S61335 | hemoglobin recepto |
| 3          | 786.5  | 18.7        | 810    | 2 A81965 | hemoglobin-naptog  |
| 4          | 759    | 18.1        | 1084   | 2 B64088 | hemoglobin-binding |
| 5          | 714.5  | 17.0        | 953    | 2 B64083 | hemoglobin-binding |
| 6          | 706.5  | 16.8        | 764    | 2 H83055 | probable outer mem |
| 7          | 685.5  | 16.3        | 944    | 2 C81798 | lactoferrin bindin |
| 8          | 681.5  | 15.8        | 943    | 2 G81070 | lactoferrin-bindin |
| 9          | 649    | 15.5        | 698    | 2 B82443 | heme transport pro |
| 10         | 641    | 15.3        | 940    | 2 S49087 | lactoferrin bindin |
| 11         | 640    | 15.2        | 931    | 2 S66574 | transferrin-bindin |
| 12         | 597.5  | 14.2        | 911    | 2 S70911 | transferrin-bindin |
| 13         | 590    | 14.0        | 914    | 2 S70906 | transferrin-bindin |
| 14         | 588    | 14.0        | 910    | 2 C81832 | transferrin-bindin |
| 15         | 587.5  | 14.0        | 911    | 2 JN0821 | transferrin-bindin |
| 16         | 582    | 13.9        | 908    | 2 JN0819 | transferrin-bindin |
| 17         | 577    | 13.7        | 912    | 2 S70901 | transferrin-bindin |
| 18         | 576.5  | 13.7        | 915    | 2 F81196 | transferrin-bindin |
| 19         | 575    | 13.7        | 912    | 2 C64107 | transferrin-bindin |
| 20         | 562.5  | 13.4        | 915    | 2 A43335 | transferrin-bindin |
| 21         | 542    | 12.9        | 877    | 2 AC2211 | heme transport pro |
| 22         | 521.5  | 12.4        | 807    | 2 AC2075 | outer membrane hem |
| 23         | 481    | 11.5        | 758    | 2 I64084 | hemoglobin recepto |
| 24         | 460.5  | 11.0        | 733    | 2 A87521 | TonB-dependent rec |
| 25         | 447    | 10.6        | 766    | 2 A82857 | outer membrane hem |
| 26         | 447    | 10.6        | 766    | 2 D97521 | probable outer mem |
| 27         | 443    | 10.5        | 518    | 2 A64130 | hypothetical prote |
| 28         | 345.5  | 8.2         | 659    | 2 G81009 | colicin I receptor |
| 29         | 341.5  | 8.1         | 659    | 2 A85854 | hypothetical prote |

|    |       |     |     |          |                     |
|----|-------|-----|-----|----------|---------------------|
| 30 | 341   | 8.1 | 830 | 2 AH0477 | probable TonB depe  |
| 31 | 337.5 | 8.0 | 663 | 1 QREIC  | colicin I receptor  |
| 32 | 336.5 | 8.0 | 663 | 2 AG0782 | colicin I receptor  |
| 33 | 323   | 7.7 | 725 | 2 A57148 | outer membrane pro  |
| 34 | 322.5 | 7.7 | 851 | 2 A83484 | outer membrane util |
| 35 | 313   | 7.5 | 723 | 2 C64058 | outer membrane pro  |
| 36 | 307   | 7.3 | 751 | 2 AC0574 | ferrienterobactin   |
| 37 | 304   | 7.2 | 713 | 2 F82506 | probable TonB syst  |
| 38 | 300   | 7.1 | 665 | 2 AD0160 | probable outer mem  |
| 39 | 299   | 7.1 | 784 | 2 D82437 | TonB receptor-rela  |
| 40 | 297.5 | 7.1 | 702 | 2 AF3038 | exogenous ferric s  |
| 41 | 297.5 | 7.1 | 707 | 2 E98247 | outer membrane pro  |
| 42 | 293   | 7.0 | 660 | 2 D91176 | heme utilization/t  |
| 43 | 293   | 7.0 | 660 | 2 E86022 | outer membrane hem  |
| 44 | 293   | 7.0 | 696 | 2 H90798 | Iha adhesin (impor  |
| 45 | 293   | 7.0 | 696 | 2 G85607 | probable receptor   |

ALIGNMENTS

RESULT 1

F81056  
hemoglobin receptor NMB1668 [imported] - Neisseria meningitidis (strain MC58 serogroup E  
C;Species: Neisseria meningitidis  
C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C;Accession: F81056  
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizzi, M.  
Science 287, 1809-1815, 2000  
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve.  
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A;Reference number: A81000; MUID:20175755; PMID:10710307  
A;Accession: F81056  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-791 <FET>  
A;Cross-references: UNIPROT:Q9JYA8; GB:AE002517; GB:AE002098; NID:G7226912; PIDN:AAF4201  
A;Experimental source: serogroup B, strain MC58  
C;Genetics:  
A;Gene: NMB1668  
C;Superfamily: ferrichrome-iron receptor; tonB-dependent receptor amino-terminal homology

|                       |                  |   |           |             |
|-----------------------|------------------|---|-----------|-------------|
| Query Match           | 92.1%            | Score 3867.5;   | DB 2;     | Length 791; |
| Best Local Similarity | 92.3%            | Pred. No. 3e-246;   |           |             |
| Matches 730;          | Conservative 20; | Mismatches 40;  | Indels 1; | Gaps 1;     |
| Qy                    | 1                | MKPLHMLPIAALVGSIFGNPVLAADEAATETTPVKAEIKERVVKDQNLNAPATVERVNLGR | 60        |             |
| Db                    | 1                | MKPLQMLPIAALVGSIFGNPVLAADEAATETTPVKAEIKAVRVKQRPAAVERVNLNR     | 60        |             |
| Qy                    | 61               | IQQEMIRDNKDVAARYSTDVGLSDSGHKGFAVRGVEGNRVGVSIDGVSLPDSSENSLYA   | 120       |             |
| Db                    | 61               | IKQEMIRDNKDVAARYSTDVGLSDSGHKGFAVRGVEGNRVGVSIDGVNLPDSSENSLYA   | 120       |             |
| Qy                    | 121              | RYGNFNSRLSIDPELVNRIEIAKAGADSFNTGSGALGGGVNYQTLOGHLLDLLDDRQFGVM | 180       |             |
| Db                    | 121              | RYGNFNSRLSIDPELVNRIEIVKGADSFNTGSGALGGGVNYQTLOGRLLDLLDDRQFGVM  | 180       |             |
| Qy                    | 181              | MKNGYSRRNREWTNTLFGVSNDRVDAALYSQRRGHETESAGERGYFVGAGSGALIRG     | 240       |             |
| Db                    | 181              | MKNGYSTNRNREWTNTLFGVSNDRVDAALYSQRRGHETESAGRGYAVGEGSGANIRG     | 240       |             |
| Qy                    | 241              | SSRGIPDPSKHKHNFGLKIAYQINDKHRIGPSFNGQQGHNYTTEESYNLTASSWREADD   | 300       |             |
| Db                    | 241              | SARGIPDSSKHKHNFGLKIAYQINDNHRIGASLNGQQGHNYTVEESYNLTASSWREADD   | 300       |             |
| Qy                    | 301              | VNRRRNANLFYEWTPDPSNWLSSLKADPDYQTTKVAAVNNKGSPPDTYSTWTRTNQKDL   | 360       |             |
| Db                    | 301              | VNRRRNANLFYEWTPDPSNWLSSLKADPDYQTTKVAAVNNKGSPPDYSTWTRTNQKDL    | 360       |             |
| Qy                    | 361              | NIYNSMDTFRKPTFLRMDSQPLQL-CGQHRLSILKTFASRRPENLRDDDYFSEVRSRT    | 419       |             |

Db 541 THTLSQGRDGTDLNLIYQSNRYNRLSEBQNLTVSGTGGCTEEDAYYYRCDPYKEKL 600  
QY 600 EWOMONIDKARIRGLTGLRLNLTWVAVFVPEGWKLFGLGVAKSKLSDGNSLLSTQPPK 659  
Db 601 DMQKNIDKARIRGLTGLRLNLTWVAVFVPEGWKLFGLGVAKSKLSDGNSLLSTQPPK 660  
QY 660 VIAGDYSPSEKMGVFSRLTYLGAKKAKDAQYTVYENKRGCTPLQKKVKDYPWLKNSAY 719  
Db 661 VIAGDYSPSEKMGVFSRLTYLGAKKAKDAQYTVYENKRGCTPLQKKVKDYPWLKNSAY 720  
QY 720 VEDMGYFKLAKNLTIRAGVYVFNRYTWTWDSLRGLYSYSTTNAVDRDCKGLDRYRAG 779  
Db 721 VEDMGYFKLAKNLTIRAGVYVFNRYTWTWDSLRGLYSYSTTNAVDRDCKGLDRYRAG 780  
QY 780 RNYAVSLDWKF 790  
Db 781 RNYAVSLEWKF 791

RESULT 10  
US-08-537-361E-4  
; Sequence 4, Application US/08537361E  
; Patent No. 6121037  
; GENERAL INFORMATION:  
; APPLICANT: Stojiljkovic, Igor  
; APPLICANT: So, Magdalene  
; APPLICANT: Hwa, Vivian  
; APPLICANT: Heffron, Fred  
; APPLICANT: Nasif, Xavier  
; TITLE OF INVENTION: No. 6121037el Bacterial Hemoglobin Receptor  
; TITLE OF INVENTION: Genes and Uses  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive, 32nd Floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/537,361E  
; FILING DATE: 02-OCT-1995  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 6121037han, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 94,784-A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-0002  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 791 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-537-361E-4

Query Match 87.3%; Score 3665.5; DB 3; Length 791;  
Best Local Similarity 86.9%; Pred. No. 0;  
Matches 687; Conservative 44; Mismatches 59; Indels 1; Gaps 1;  
QY 1 MKPLHMLPIAALVGSIFGNPVLAADEATETTPVKAIEKVRVKDQLNAPATVERVNLGR 60  
Db 1 MKPLQWPIAALLGSIIFGNVFADEATETTPVKAIEKAVRVKQGNAPAAVERVNLNR 60  
QY 61 IQQEMIRDNKDLVRYSTDVGLSDSRHQKGFPAIRGVGDRVGVSDGVNLPDSEENSLYA 120

Db 61 IKQEMIRDNKDLVRYSTDVGLSDSRHQKGFPAIRGVGDRVGVSDGVNLPDSEENSLYA 120  
QY 121 RYGNFNSRLSIDPELVNIEIAKADGFNTGSGALGGGVNYQTLOGHLLDLDROFGVM 180  
Db 121 RYGNFNSRLSIDPELVNIDIVKADSFNTGSGALGGGVNYQTLOGRDLILLPERQFGVM 180  
QY 181 MKNGYSSRNREWTNLTGFGVSNDRVDAALLYSQRRGHETESAGERGYPVEGAGSGAIRG 240  
Db 181 MKNGYSTNRREWTNLTGFGVSNDRVDAALLYSQRRGHETESAGERGYPVEGAGSGAIRG 240  
QY 241 SSRGIPDPSKHYHNFGLKIAIYQINDKHRIGPSFNGQOQHNYTIEESNYLTASSHREADD 300  
Db 241 SARGIPDPSQHKYHFLGKIAIYQINDKHRI GASLNGOQOQHNYTIEESNYLTASSYREADD 300  
QY 301 VNRRNANLFYEWTPDPSNLSLKADFDYQTTKVAANVKNKGSFPPTDYSTWTNRYNOKOLE 360  
Db 301 VNRRNTNLFYEWTPDPSNLSLVKADVDYQTKVSAVNYKGSFPPTNYTWTWEYHKKEVG 360  
QY 361 NIYNRSMDTRFKRFTLRMDSQPLQL-GGQHRLSLTKTFASRRREFENLRDDYFYSERSVRT 419  
Db 361 EIYNRSMDTTFKRI TLRMDSHPLQLGGGRHRLSFKTFAGQRDFENLRDDYFYSGRVVRT 420  
QY 420 TSSIOHPVKTNYGFSLSDOIQWINDVFSRADIRYDHTKMTPOELNAEACHADCKTPPAAN 479  
Db 421 TNSIOHPVKTNYGFSLSDOIQWINDVFSRADIRYDHTKMTPOELNADCHADCKTPPAAN 480  
QY 480 TYKMGSGFVGLAAQLNQAHVGYDITSGYRVVPNASEVFTYNHSGSNWLPNPKAERST 539  
Db 481 TYKMGSGFVGLAAQLSQTWRGLGYDVTSGFRVPNASEVFTYNHSGSTWKFENPKAERST 540  
QY 540 THTLSQGRSEKGTLDANLYQNNYRNFLEBQKLTTSBGVGTQMNYYGMCNPSYSEKP 599  
Db 541 THTLSQGRGDKGTLNLYQSNRYNRLSEBQNLTVSGTGGCTEEDAYYYRCDPYKEKL 600  
QY 600 EWOMONIDKARIRGLTGLRLNLTWVAVFVPEGWKLFGLGVAKSKLSDGNSLLSTQPPK 659  
Db 601 DMQKNIDKARIRGLTGLRLNLTWVAVFVPEGWKLFGLGVAKSKLSDGNSLLSTQPPK 660  
QY 660 VIAGDYSPSEKMGVFSRLTYLGAKKAKDAQYTVYENKRGCTPLQKKVKDYPWLKNSAY 719  
Db 661 VIAGDYSPSEKMGVFSRLTYLGAKKAKDAQYTVYENKRGCTPLQKKVKDYPWLKNSAY 720  
QY 720 VEDMGYFKLAKNLTIRAGVYVFNRYTWTWDSLRGLYSYSTTNAVDRDCKGLDRYRAG 779  
Db 721 VEDMGYFKLAKNLTIRAGVYVFNRYTWTWDSLRGLYSYSTTNAVDRDCKGLDRYRAG 780  
QY 780 RNYAVSLDWKF 790  
Db 781 RNYAVSLEWKF 791

RESULT 11  
US-09-809-665A-32  
; Sequence 32, Application US/09809665A  
; Patent No. 6790950  
; GENERAL INFORMATION:  
; APPLICANT: Lowery E., David, et.al.  
; TITLE OF INVENTION: Anti-Bacterial Vaccine Compositions  
; FILE REFERENCE: 28341/00435  
; CURRENT APPLICATION NUMBER: US/09/809,665A  
; CURRENT FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 60/153,453  
; PRIOR FILING DATE: 1999-09-10  
; PRIOR APPLICATION NUMBER: 60/128,689  
; PRIOR FILING DATE: 1999-04-09  
; PRIOR APPLICATION NUMBER: 09/545,199  
; NUMBER OF SEQ ID NOS: 197  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 32  
; LENGTH: 967  
; TYPE: PRT  
; ORGANISM: Pasteurella multocida

Best Local Similarity 90.9%; Pred. No. 0;  
Matches 720; Conservative 25; Mismatches 45; Indels 2; Gaps 2;  
Qy 1 MKPLHMLPIAALVGSIFGNPVLAADEAATETTPVKAIEKEVRVKDQLNAPATVERVNLGR 60  
Db 1 MKPLQMLPIAALVGSIFGNPVLAADEAATETTPVKAIEKAVRGQRNAPAAVERVNLNR 60  
Qy 61 IQQEMIRNDKLVRYSTDVGLSDSGRHKQFAVRGVEGNRVGVSDIGVSLPDSSENSLYA 120  
Db 61 IQQEMIRNDKLVRYSTDVGLSDSGRHKQFAVRGVEGNRVGVSDIGVSLPDSSENSLYA 120  
Qy 121 RYGNFNSRLSIDPELVNRIEIAKADSFNTGSGALGGVNYQTLQGHDLDDRRQFGVM 180  
Db 121 RYGNFNSRLSIDPELVNRIEIAKADSFNTGSGALGGVNYQTLQGHDLDDRRQFGVM 180  
Qy 181 MKNGYSRNREWTNTLFGVSNDRVDAALYSORRGHETESAGRGYPVEGAGSGAIRG 240  
Db 181 MKNGYSRNREWTNTLFGVSNDRVDAALYSORRGHETESAGRGYPVEGAGSGAIRG 240  
Qy 241 SSRGIPDPSKHYNFLGKIAYQINDKHRIQSPFNGOQHNYTIEESYNLTASSWRADD 300  
Db 241 SARGIPDPSKHYNFLGKIAYQINDKHRIQSPFNGOQHNYTIEESYNLTASSWRADD 300  
Qy 301 VNRNRNANLFYEWTPDSNWLSSKADFDYQTKVAANVNNKGSFP-TDYSTWTRNYNQKDL 359  
Db 301 VNRNRNANLFYEWTPDSNWLSSKADFDYQTKVAANVNNKGSFP-TDYSTWTRNYNQKDL 359  
Qy 360 ENIYNSRMDTRFKPTLRMDSDQIOWNDVFSRADIRYDHTKMTPOELNAECHACDKTPPAA 418  
Db 360 ENIYNSRMDTRFKPTLRMDSDQIOWNDVFSRADIRYDHTKMTPOELNAECHACDKTPPAA 418  
Qy 419 TTSSIQHPVKTNTYGFSLSDQIOWNDVFSRADIRYDHTKMTPOELNAECHACDKTPPAA 478  
Db 419 TTSSIQHPVKTNTYGFSLSDQIOWNDVFSRADIRYDHTKMTPOELNAECHACDKTPPAA 478  
Qy 479 NTYKWSGFGVGLAAQLNOAHVGVYDITSGYRVPNASEVYFTYHNSGNGWLPNPNLKAERS 538  
Db 479 NTYKWSGFGVGLAAQLNOAHVGVYDITSGYRVPNASEVYFTYHNSGNGWLPNPNLKAERS 538  
Qy 539 TTHTLSLQGRSEKGTLDANLYQNNYRNFLSEQKLTTSQDVGCTOMNYYYGMCNPNYSEK 598  
Db 539 TTHTLSLQGRSEKGTLDANLYQNNYRNFLSEQKLTTSQDVGCTOMNYYYGMCNPNYSEK 598  
Qy 541 TTHTLSLQGRSEKGTLDANLYQNNYRNFLSEQKLTTSQDVGCTOMNYYYGMCNPNYSEK 600  
Db 541 TTHTLSLQGRSEKGTLDANLYQNNYRNFLSEQKLTTSQDVGCTOMNYYYGMCNPNYSEK 600  
Qy 599 PEWQMNIDKARIRGELTGRNLNTKVASFVPEGWKLFGLGKSLGNSLLSTOPP 658  
Db 599 PEWQMNIDKARIRGELTGRNLNTKVASFVPEGWKLFGLGKSLGNSLLSTOPP 658  
Qy 601 LEWQMNIDKARIRGELTGRNLNTKVASFVPEGWKLFGLGKSLGNSLLSTOPP 660  
Db 601 LEWQMNIDKARIRGELTGRNLNTKVASFVPEGWKLFGLGKSLGNSLLSTOPP 660  
Qy 659 KVIAGVDYSPSEKGVFSRLTYLGAKKAKDAQYTVYENKGRGTPLOKKVKDY PWNKSA 718  
Db 659 KVIAGVDYSPSEKGVFSRLTYLGAKKAKDAQYTVYENKGRGTPLOKKVKDY PWNKSA 718  
Qy 719 YVFDYGFYKLAKNLTLRAGVYNNYRNKTYTWDLSRLGLYSSTINAVDRDCKGLDRYRAS 778  
Db 719 YVFDYGFYKLAKNLTLRAGVYNNYRNKTYTWDLSRLGLYSSTINAVDRDCKGLDRYRAS 778  
Qy 721 YVFDYGFYKLAKNLTLRAGVYNNYRNKTYTWDLSRLGLYSSTINAVDRDCKGLDRYRAS 780  
Db 721 YVFDYGFYKLAKNLTLRAGVYNNYRNKTYTWDLSRLGLYSSTINAVDRDCKGLDRYRAS 780  
Qy 779 GRNAVSLDWKF 790  
Db 779 GRNAVSLDWKF 790  
Qy 781 SRNAVSLDWKF 792  
Db 781 SRNAVSLDWKF 792

RESULT 9  
US-08-817-707-4  
; Sequence 4, Application US/08817707  
; Patent No. 6277382  
; GENERAL INFORMATION:  
; APPLICANT: Stojiljkovic, Igor  
; APPLICANT: So, Magdalene  
; APPLICANT: Hwa, Vivian  
; APPLICANT: Heffron, Fred  
; APPLICANT: Nassif, Xavier  
; TITLE OF INVENTION: No. 6277382el Bacterial Hemoglobin Receptor  
; TITLE OF INVENTION: Genes and Uses  
; NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:  
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
STREET: 300 South Wacker Drive, 32nd Floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817,707  
FILING DATE: 19-AUG-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6277382nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 94,784-J  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 791 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-817-707-4

Query Match 87.3%; Score 3668.5; DB 3; Length 791;  
Best Local Similarity 87.0%; Pred. No. 0;  
Matches 688; Conservative 43; Mismatches 59; Indels 1; Gaps 1;  
Qy 1 MKPLHMLPIAALVGSIFGNPVLAADEAATETTPVKAIEKEVRVKDQLNAPATVERVNLGR 60  
Db 1 MKPLQMLPIAALVGSIFGNPVLAADEAATETTPVKAIEKAVRGQRNAPAAVERVNLNR 60  
Qy 61 IQQEMIRNDKLVRYSTDVGLSDSGRHKQFAVRGVEGNRVGVSDIGVSLPDSSENSLYA 120  
Db 61 IQQEMIRNDKLVRYSTDVGLSDSGRHKQFAVRGVEGNRVGVSDIGVSLPDSSENSLYA 120  
Qy 121 RYGNFNSRLSIDPELVNRIEIAKADSFNTGSGALGGVNYQTLQGHDLDDRRQFGVM 180  
Db 121 RYGNFNSRLSIDPELVNRIEIAKADSFNTGSGALGGVNYQTLQGHDLDDRRQFGVM 180  
Qy 181 MKNGYSRNREWTNTLFGVSNDRVDAALYSORRGHETESAGRGYPVEGAGSGAIRG 240  
Db 181 MKNGYSRNREWTNTLFGVSNDRVDAALYSORRGHETESAGRGYPVEGAGSGAIRG 240  
Qy 241 SSRGIPDPSKHYNFLGKIAYQINDKHRIQSPFNGOQHNYTIEESYNLTASSWRADD 300  
Db 241 SARGIPDPSKHYNFLGKIAYQINDKHRIQSPFNGOQHNYTIEESYNLTASSWRADD 300  
Qy 301 VNRNRNANLFYEWTPDSNWLSSKADFDYQTKVAANVNNKGSFP-TDYSTWTRNYNQKDL 360  
Db 301 VNRNRNANLFYEWTPDSNWLSSKADFDYQTKVAANVNNKGSFP-TDYSTWTRNYNQKDL 360  
Qy 361 NIYNSRMDTRFKPTLRMDSDQIOWNDVFSRADIRYDHTKMTPOELNAECHACDKTPPAA 419  
Db 361 NIYNSRMDTRFKPTLRMDSDQIOWNDVFSRADIRYDHTKMTPOELNAECHACDKTPPAA 419  
Qy 420 TTSSIQHPVKTNTYGFSLSDQIOWNDVFSRADIRYDHTKMTPOELNAECHACDKTPPAA 479  
Db 420 TTSSIQHPVKTNTYGFSLSDQIOWNDVFSRADIRYDHTKMTPOELNAECHACDKTPPAA 479  
Qy 480 TYKWSGFGVGLAAQLNOAHVGVYDITSGYRVPNASEVYFTYHNSGNGWLPNPNLKAERS 539  
Db 480 TYKWSGFGVGLAAQLNOAHVGVYDITSGYRVPNASEVYFTYHNSGNGWLPNPNLKAERS 539  
Qy 540 TTHTLSLQGRSEKGTLDANLYQNNYRNFLSEQKLTTSQDVGCTOMNYYYGMCNPNYSEK 599  
Db 540 TTHTLSLQGRSEKGTLDANLYQNNYRNFLSEQKLTTSQDVGCTOMNYYYGMCNPNYSEK 599

GENERAL INFORMATION:  
 APPLICANT: Stojiljkovic, Igor  
 APPLICANT: So, Magdalene  
 APPLICANT: Hwa, Vivian  
 APPLICANT: Heffron, Fred  
 APPLICANT: Nassif, Xavier  
 TITLE OF INVENTION: No. 6121037el Bacterial Hemoglobin Receptor  
 TITLE OF INVENTION: Genes and Uses  
 NUMBER OF SEQUENCES: 14  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff  
 STREET: 300 South Wacker Drive, 32nd Floor  
 CITY: Chicago  
 STATE: Illinois  
 COUNTRY: USA  
 ZIP: 60606  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/537,361E  
 FILING DATE: 02-OCT-1995  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: No. 6121037nan, Kevin E  
 REGISTRATION NUMBER: 35,303  
 REFERENCE/DOCKET NUMBER: 94,784-A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312-913-0001  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 792 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-537-361E-6

Query Match 90.5%; Score 3806; DB 3; Length 792;  
 Best Local Similarity 91.2%; Pred. No. 0;  
 Matches 722; Conservative 24; Mismatches 44; Indels 2; Gaps 2;

|    |     |  |     |
|----|-----|--|-----|
| QY | 1   | MPKPLHMLPAAALVGSIFGNPVLAADEAATETTPVKAIEKVRVKDQLNAPATVERVNLGR | 60  |
| DB | 1   | MPKPLQMLPAAALVGSIFGNPVLAADEAATETTPVKAIEKVRVKDQLNAPATVERVNLGR | 60  |
| QY | 61  | IQEIMRDNDKLVRYSTDVGLSDSGRHKQGFVAVRVEGNRNVGSIDGVSPLPDSEENSLYA | 120 |
| DB | 61  | IQEIMRDNDKLVRYSTDVGLSDSGRHKQGFVAVRVEGNRNVGSIDGVSPLPDSEENSLYA | 120 |
| QY | 121 | RYGNFNSSRLSDPELVRIETAKGADSFNTGSGALGGGVNYQTLOCHDLLLDDRQFGVM   | 180 |
| DB | 121 | RYGNFNSSRLSDPELVRIETAKGADSFNTGSGALGGGVNYQTLOCHDLLLDDRQFGVM   | 180 |
| QY | 181 | MNGVSSRNKRWNTLGFVGSNDVDAALLYSORRHETESAGERYPVEGAGSGAIIRG      | 240 |
| DB | 181 | MNGVSSRNKRWNTLGFVGSNDVDAALLYSORRHETESAGERYPVEGAGSGAIIRG      | 240 |
| QY | 241 | SRGIPDPKSKHYNPLGKIAYQINDKHRIGPSFNQOQHNYTIESYNTLTASSWREADD    | 300 |
| DB | 241 | SARGIPDPKSKHYNPLGKIAYQINDKHRIGPSFNQOQHNYTIESYNTLTASSWREADD   | 300 |
| QY | 301 | VNRRNRNLFYEWTPDSNWLSSKADPDYQTTKVAIVNKGSPF-TDYSTWNTNPKDL      | 359 |
| DB | 301 | VNRRNRNLFYEWTPDSNWLSSKADPDYQTTKVAIVNKGSPF-TDYSTWNTNPKDL      | 359 |
| QY | 360 | ENLYNRSMDTRFKRFTLRMDSPQL-L-GGCHRLSLKTFASRRFENLNRDDYFFSERVR   | 418 |
| DB | 361 | DELYNRSMDTRFKRFTLRMDSPQL-L-GGCHRLSLKTFASRRFENLNRDDYFFSERVR   | 418 |
| QY | 419 | TTSSIOHPVKTTNYGFSLSDDQIQWNVFSSRAGIRYDHTKMTPOELNAEACHADKTPPAA | 478 |

|    |     |   |     |
|----|-----|---|-----|
| DB | 421 | TTSSIOHPVKTTNYGFSLSDDQIQWNVFSSRAGIRYDHTKMTPOELNAEACHADKTPPAA  | 480 |
| QY | 479 | NTYKWSGFGVGLAAQLNQAHWVGYDITSGYRVPNAASEVYFTYNNHSGNWLPNPNLKAERS | 538 |
| DB | 481 | NTYKWSGFGVGLAAQLNQAHWVGYDITSGYRVPNAASEVYFTYNNHSGNWLPNPNLKAERS | 540 |
| QY | 539 | TTHTLSLQSRSEKGTLDANLYQNNYRNFLSEBQKLTITSGDVGCTQWNYYYGMCNPPYSEK | 598 |
| DB | 541 | TTHTLSLQSRSEKGTLDANLYQNNYRNFLSEBQKLTITSGDVGCTQWNYYYGMCNPPYSEK | 600 |
| QY | 599 | PEWQONIDKARIRGLELTGRNLNVTKVASFVPEGHKLFGLSLGYAKSKLSGDNLSLSTQPP | 658 |
| DB | 601 | LEWQONIDKARIRGLELTGRNLNVTKVASFVPEGHKLFGLSLGYAKSKLSGDNLSLSTQPP | 660 |
| QY | 659 | KVIAGVDYESPEKGVFSRLTYLCAKKAQADQYTYENKGRGTPLOKKVKDYPWLNKSA     | 718 |
| DB | 661 | KVIAGIDYESPEKGVFSRLTYLCAKKAQADQYTYENKGRGTPLOKKVKDYPWLNKSA     | 720 |
| QY | 719 | YVFDMYGFYKLAKNLTLAGVYVNFNRKYTTWDSLRGLYSYSTTNNAVDRDGKGLDRYRAS  | 778 |
| DB | 721 | YVFDMYGFYKPVKNLTLAGVYVNFNRKYTTWDSLRGLYSYSTTNNAVDRDGKGLDRYRAP  | 780 |
| QY | 779 | GRNYAVSLDWKF 790  |     |
| DB | 781 | SRNYAVSLEWKF 792  |     |

## RESULT 8

US-08-537-361E-2  
 ; Sequence 2, Application US/08537361E  
 ; Patent No. 6121037  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Stojiljkovic, Igor  
 ; APPLICANT: So, Magdalene  
 ; APPLICANT: Hwa, Vivian  
 ; APPLICANT: Heffron, Fred  
 ; APPLICANT: Nassif, Xavier  
 ; TITLE OF INVENTION: No. 6121037el Bacterial Hemoglobin Receptor  
 ; TITLE OF INVENTION: Genes and Uses  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff  
 ; STREET: 300 South Wacker Drive, 32nd Floor  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/537,361E  
 ; FILING DATE: 02-OCT-1995  
 ; CLASSIFICATION: 536  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: No. 6121037nan, Kevin E  
 ; REGISTRATION NUMBER: 35,303  
 ; REFERENCE/DOCKET NUMBER: 94,784-A  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312-913-0001  
 ; TELEFAX: 312-913-0002  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 792 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-537-361E-2

Query Match 90.5%; Score 3799; DB 3; Length 792;

Qy 301 VNRNRNANLFYEWTPDSNWLSSKADFDYQTKVAANVNGSGFP-TDYSTWTRNYNQKDL 359  
Db 301 VNRNRNANLFYEWTPDSNWLSSKADFDYQTKVAANVNGSGFP-TDYSTWTRNYNQKDL 360  
Qy 360 ENIYNRSMDFRKFPTLRMDSQPLQ-LGGHRLSLKTFASRRFENLNRDDYYFSSVR 418  
Db 361 DEIYNRSMDFRKFPTLRMDSQPLQ-LGGHRLSLKTFASRRFENLNRDDYYFSSVR 420  
Qy 419 TTSIQHPVKTNTYGFSLSDQIQWNVFSSRADIRYDHTKMTPOELNAEHCADKTPPAA 478  
Db 421 TTSIQHPVKTNTYGFSLSDQIQWNVFSSRADIRYDHTKMTPOELNAEHCADKTPPAA 480  
Qy 479 NTYKWSGFGVGLAAQLNAQAHVGYDITSGYRVPNASEVYFTYHNSGNGWLPNPNLKAERS 538  
Db 481 NTYKWSGFGVGLAAQLNAQAHVGYDITSGYRVPNASEVYFTYHNSGNGWLPNPNLKAERT 540  
Qy 539 THTLSLQGRSEKGTLDANLYQNNYRNFPLSBEQKLTTSQDVCGTQMMYYGMCNPNYSEK 598  
Db 541 THTLSLQGRSEKGTLDANLYQNNYRNFPLSBEQKLTTSQDVCGTQMMYYGMCNPNYSEK 600  
Qy 599 PEMQONIDKARIRGLELTGRLNVTYKASFPVEGKMLFGSLGYAKSLSGDNSLLSTOPP 658  
Db 601 PEMQONIDKARIRGLELTGRLNVTYKASFPVEGKMLFGSLGYAKSLSGDNSLLSTOPP 660  
Qy 659 KVIAGVDYSESPSEKGVFSRLTYLGAKKAKADQYTVYENKGRGTPLOKKVKDYPWLKNSA 718  
Db 661 KVIAGVDYSESPSEKGVFSRLTYLGAKKAKADQYTVYENKGRGTPLOKKVKDYPWLKNSA 720  
Qy 719 YVFDYGFYKAKNLTIRAGYVNFNRKYTTWDSLRGLYSYSTTNAVDRDGLDRYRAS 778  
Db 721 YVFDYGFYKAKNLTIRAGYVNFNRKYTTWDSLRGLYSYSTTNAVDRDGLDRYRAS 780  
Qy 779 GRNVAVSLDWKF 790  
Db 781 SRNVAVSLDWKF 792

## RESULT 6

US-08-817-707-6  
; Sequence 6, Application US/08817707  
; Patent No. 6277382  
; GENERAL INFORMATION:  
; APPLICANT: Stojiljkovic, Igor  
; APPLICANT: So, Magdalene  
; APPLICANT: Hwa, Vivian  
; APPLICANT: Heffron, Fred  
; APPLICANT: Nassif, Xavier  
; TITLE OF INVENTION: No. 6277382el Bacterial Hemoglobin Receptor  
; TITLE OF INVENTION: Genes and Uses  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive, 32nd Floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/817,707  
; FILING DATE: 19-AUG-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 6277382nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 94,784-J  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-0002

; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 790 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-817-707-6

Query Match 90.7%; Score 3810; DB 3; Length 790;  
Best Local Similarity 90.9%; Pred. No. 0;  
Matches 719; Conservative 26; Mismatches 44; Indels 2; Gaps 2;

Qy 1 MKPLHMLPIAALVSGIFGNFVLADEAAETTPVKAEIKVVRVQDLNAPATVVRVNLGR 60  
Db 1 MKPLQMLPIAALVSGIFGNFVLADEAAETTPVKAEIKAVRVKQGNAPAAVERVNLNR 60  
Qy 61 IQEMIRDNKDLVRYSTDVGLSDSGRHKQGFVAVRVEGNRVGVISIDGVSIPDSENSLYA 120  
Db 61 IQEMIRDNKDLVRYSTDVGLSDSGRHKQGFVAVRVEGNRVGVISIDGVSIPDSENSLYA 120  
Qy 121 RYGNFNSRSLSDPELVNRNIEIAKADSPNTGSGALGGVNYQTLQGHDLDDDRQGVN 180  
Db 121 RYGNFNSRSLSDPELVNRNIEIVKADSPNTGSGALGGVNYQTLQGRDLDLDDDRQGVN 180  
Qy 181 MKNGYSRRNREWTNTLGFVSNDRVDAALYSQRRGHETESAGRGYPVEGAGSIAIRG 240  
Db 181 MKNGYSRRNREWTNTLGFVSNDRVDAALYSQRRGHETESAGRGYPVEGAGKETNIRG 240  
Qy 241 SSRGIPDPSKHVNFHFLGIAYQINDKRIITGSPNGOQGHNYTIERSYNLTASSWRBADD 300  
Db 241 SARGIPDPSKHVNFHFLGIAYQINDNHRICASLNGOQGHNYTVEESYNLTASSWRBADD 300  
Qy 301 VNRERNANLFYEWTPDSNWLSSKADFDYQTKVAANVNGSGFPDYSTWTRNYNQKDL 360  
Db 301 VNRERNANLFYEWTPDSNWLSSKADFDYQTKVAANVNGSGFPDYSTWTRNYNQKDL 360  
Qy 361 NIYNRSMDFRKFPTLRMDSQPLQ-LGGHRLSLKTFASRRFENLNRDDYYFSSVR 419  
Db 361 NIYNRSMDFRKFPTLRMDSQPLQ-LGGHRLSLKTFASRRFENLNRDDYYFSSVR 419  
Qy 420 TTSIQHPVKTNTYGFSLSDQIQWNVFSSRADIRYDHTKMTPOELNAEHCADKTPPAA 479  
Db 420 TTSIQHPVKTNTYGFSLSDQIQWNVFSSRADIRYDHTKMTPOELNAEHCADKTPPAA 479  
Qy 480 TYKWSGFGVGLAAQLNAQAHVGYDITSGYRVPNASEVYFTYHNSGNGWLPNPNLKAERS 539  
Db 480 TYKWSGFGVGLAAQLNAQAHVGYDITSGYRVPNASEVYFTYHNSGNGWLPNPNLKAERS 539  
Qy 540 THTLSLQGRSEKGTLDANLYQNNYRNFPLSBEQKLTTSQDVCGTQMMYYGMCNPNYSEK 599  
Db 540 THTLSLQGRSEKGTLDANLYQNNYRNFPLSBEQKLTTSQDVCGTQMMYYGMCNPNYSEK 599  
Qy 600 EWMQONIDKARIRGLELTGRLNVTYKASFPVEGKMLFGSLGYAKSLSGDNSLLSTOPP 659  
Db 600 EWMQONIDKARIRGLELTGRLNVTYKASFPVEGKMLFGSLGYAKSLSGDNSLLSTOPP 659  
Qy 660 VIAGVDYSESPSEKGVFSRLTYLGAKKAKADQYTVYENKGRGTPLOKKVKDYPWLKNSA 719  
Db 660 VIAGVDYSESPSEKGVFSRLTYLGAKKAKADQYTVYENKGRGTPLOKKVKDYPWLKNSA 719  
Qy 720 VFDYGFYKAKNLTIRAGYVNFNRKYTTWDSLRGLYSYSTTNAVDRDGLDRYRAS 779  
Db 720 VFDYGFYKAKNLTIRAGYVNFNRKYTTWDSLRGLYSYSTTNAVDRDGLDRYRAS 779  
Qy 780 RNYAVSLDWKF 790  
Db 780 RNYAVSLDWKF 790

## RESULT 7

US-08-537-361E-6  
; Sequence 6, Application US/08537361E  
; Patent No. 6121037

FILING DATE: 19-AUG-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: NO. 6277382han, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 94,784-J  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 792 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-817-707-2

Query Match 90.8%; Score 3814; DB 3; Length 792;  
Best Local Similarity 91.3%; Pred. No. 0;  
Matches 723; Conservative 24; Mismatches 43; Indels 2; Gaps 2;

QY 1 MKPLHMLPTAALVGSIFGNPVLAADEAATETTPVKAIEKVRVKQOLNAPATVERVNLGR 60  
DB 1 MKPLQMLPTAALVGSIFGNPVLAADEAATETTPVKAIEKVRVKQOLNAPATVERVNLNR 60

QY 61 IQQEMIRNDKDLVRYSTDVGLSDSGRHKQGFVAVRGVEGNRVGVSIDGVSIPDSEENSLYA 120  
DB 61 IQQEMIRNDKDLVRYSTDVGLSDSGRHKQGFVAVRGVEGNRVGVSIDGVSIPDSEENSLYA 120

QY 121 RYGNFNSSRLSIDPELVARNIEAKGADSFNTGSGALGGGVNYQTLLQGHDLDDRRQFGVM 180  
DB 121 RYGNFNSSRLSIDPELVARNIDIVKGADSFNTGSGALGGGVNYQTLLQGRDLLLPERQFGVM 180

QY 181 MKNGYSSRNREWTNTLFGVSNDRVDAALLYSQRRGHETESAGRGYPVEGAGSGAIRG 240  
DB 181 MKNGYSSRNREWTNTLFGVSNDRVDAALLYSQRRGHETESAGRGYPVEGAGSGAIRG 240

QY 241 SARGIPDPSPQHKYHFLGKIAYQINDKIRIGPSFNGQGHNTTIESYNTLTASSWREADD 300  
DB 241 SARGIPDPSPQHKYHFLGKIAYQINDKIRIGPSFNGQGHNTTIESYNTLTASSWREADD 300

QY 301 VNRNRNANLFYEWTPDSNLSLKADFDYQTTKVAANNKSGFP-TDYSWTTRNTYQKDL 359  
DB 301 VNRNRNANLFYEWTPDSNLSMKADVDYQTKVSAVNYKGSFPFIEDSTLTNTYQKDL 360

QY 360 ENIYRSMOTRPFRTLRWDSPQL-GGQRLSLKTPASRRPEFENLNRDDYFSEVRVSR 418  
DB 361 DEIYRSMOTRPFRTLRWDSPQL-GGQRLSLKTPASRRPEFENLNRDDYFSEVRVSR 420

QY 419 TTSSIQHPVKTNYGFSLSQIQWINDVFSRRADIRYDHTKMTPOELNAEACHADKTPPAA 478  
DB 421 TTSSIQHPVKTNYGFSLSQIQWINDVFSRRADIRYDHTKMTPOELNAEACHADKTPPAA 480

QY 479 NTYKWSGFGVGLAAQLNQAHWGVYDITSGYRVNASEVFTYHNGSGNWLNPNLKAERS 538  
DB 481 NTYKWSGFGVGLAAQLNQAHWGVYDITSGYRVNASEVFTYHNGSGNWLNPNLKAERS 540

QY 539 THTLSLQGRSEKGTLDANLYQNNYNFLSEBOKLTTSQDVGCTQNNYHGCSPYSEK 598  
DB 541 THTLSLQGRSEKGTLDANLYQNNYNFLSEBOKLTTSQDVGCTQNNYHGCSPYSEK 600

QY 599 PEMQONIDKARIRGLTGRNLNVTKVASFPVPGWKLFGSLGVAKSLSGDNLSTLSTOPP 658  
DB 601 PEMQONIDKARIRGLTGRNLNVTKVASFPVPGWKLFGSLGVAKSLSGDNLSTLSTOPP 660

QY 659 KVIAGVDYESPSKGVFSLTYLGAQKADQYTYENKRGTPLOKVKDYDPLINKSA 718  
DB 661 KVIAGVDYESPSKGVFSLTYLGAQKADQYTYENKRGTPLOKVKDYDPLINKSA 720

QY 719 YPDMYGFYKLANLTLRAGVNVFNRYKTTWDSLRGLYSYSTTNNAVDKRGKGLDRYAS 778  
DB 721 YPDMYGFYKLANLTLRAGVNVFNRYKTTWDSLRGLYSYSTTNNAVDKRGKGLDRYAS 780

QY 779 GRNVAVSLDWKF 790  
DB 781 SRNVAVSLDWKF 792

RESULT 5  
US-08-326-670A-2  
Sequence 2, Application US/083266670A  
Patent No. 5698438  
GENERAL INFORMATION:  
APPLICANT: Stojiljkovic, Igor  
APPLICANT: SO, Magdalene  
APPLICANT: Hwa, Vivian  
APPLICANT: Heffron, Fred  
APPLICANT: Nassif, Xavier  
TITLE OF INVENTION: A No. 5698438el Bacterial Hemoglobin Receptor Gene and Uses  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Allegretti & Witcoff, Ltd.  
STREET: 10 South Wacker Drive, Suite 3000  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/326,670A  
FILING DATE: 18 OCT 1994  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: NO. 5698438han, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 94,784  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-715-1000  
TELEFAX: 312-715-1234  
TELEX: 810-221-8317  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 792 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-326-670A-2

Query Match 90.8%; Score 3812; DB 1; Length 792;  
Best Local Similarity 91.2%; Pred. No. 0;  
Matches 722; Conservative 25; Mismatches 43; Indels 2; Gaps 2;

QY 1 MKPLHMLPTAALVGSIFGNPVLAADEAATETTPVKAIEKVRVKQOLNAPATVERVNLGR 60  
DB 1 MKPLQMLPTAALVGSIFGNPVLAADEAATETTPVKAIEKVRVKQOLNAPATVERVNLNR 60

QY 61 IQQEMIRNDKDLVRYSTDVGLSDSGRHKQGFVAVRGVEGNRVGVSIDGVSIPDSEENSLYA 120  
DB 61 IQQEMIRNDKDLVRYSTDVGLSDSGRHKQGFVAVRGVEGNRVGVSIDGVSIPDSEENSLYA 120

QY 121 RYGNFNSSRLSIDPELVARNIEAKGADSFNTGSGALGGGVNYQTLLQGHDLDDRRQFGVM 180  
DB 121 RYGNFNSSRLSIDPELVARNIDIVKGADSFNTGSGALGGGVNYQTLLQGRDLLLPERQFGVM 180

QY 181 MKNGYSSRNREWTNTLFGVSNDRVDAALLYSQRRGHETESAGRGYPVEGAGSGAIRG 240  
DB 181 MKNGYSSRNREWTNTLFGVSNDRVDAALLYSQRRGHETESAGRGYPVEGAGSGAIRG 240

QY 241 SARGIPDPSPQHKYHFLGKIAYQINDKIRIGPSFNGQGHNTTIESYNTLTASSWREADD 300  
DB 241 SARGIPDPSPQHKYHFLGKIAYQINDKIRIGPSFNGQGHNTTIESYNTLTASSWREADD 300

Qy 661 IAGVDYSPSEKMGVFSRLTYLGAKKADAQYTVYENKRGCTPLQKKVKDYPWLNKSAVV 720  
Db 661 IAGIDYSPSEKMGVFSRLTYLGAKKADAQYTVYENKRGCTPLQKKVKDYPWLNKSAVV 720  
Qy 721 FDMTGFYKLAQNLTLRAGVYVFNKRYTTWDSLRGLYSYSTTNAVDRDGKGLDRYASGR 780  
Db 721 FDMTGFYKLAQNLTLRAGVYVFNKRYTTWDSLRGLYSYSTTNAVDRDGKGLDRYASGR 780  
Qy 781 NYAVSLDWKF 790  
Db 781 NYAVSLDWKF 790  
RESULT 3  
US-08-990-470A-2  
; Sequence 2, Application US/08990470A  
; Patent No. 6123942  
; GENERAL INFORMATION:  
; APPLICANT: Stojiljkovic, Igor  
; APPLICANT: So, Magdalene  
; APPLICANT: Hwa, Vivian  
; APPLICANT: Heffron, Fred  
; APPLICANT: Nassif, Xavier  
; TITLE OF INVENTION: No. 6123942el Bacterial Hemoglobin Receptor  
; TITLE OF INVENTION: Genes and Uses  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive, 32nd Floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/990,470A  
; APPLICATION NUMBER: US/08/990,470A  
; FILING DATE: 15-DEC-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 6123942nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 94,784-K  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-0002  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 792 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-990-470A-2  
Query Match 90.8%; Score 3814; DB 3; Length 792;  
Best Local Similarity 91.3%; Pred. No. 0;  
Matches 723; Conservative 24; Mismatches 43; Indels 2; Gaps 2;  
Qy 1 MKPLHMLPIAALVGSIFGNPVLAADEAATETTPVKAIEKVRVKDQLNAPATVERVNLGR 60  
Db 1 MKPLQMLPIAALVGSIFGNPVLAADEAATETTPVKAIEKVRVKDQLNAPATVERVNLGR 60  
Qy 61 IQQEMIRNDKDLVRYSTDVGLSDSGRHKQGFVAVRGVGNRVGSDIGVSLPDSSESLIYA 120  
Db 61 IQQEMIRNDKDLVRYSTDVGLSDSGRHKQGFVAVRGVGNRVGSDIGVSLPDSSESLIYA 120  
Qy 121 RYGNFNSRLSIDPELVNRIEIAKGADSFNTGSGALGGGVNYQTLOGHDLILLDRQFGVM 180  
Db 121 RYGNFNSRLSIDPELVNRIEIAKGADSFNTGSGALGGGVNYQTLOGHDLILLDRQFGVM 180

Qy 181 MKNGYSSRNREWTNTLGFVGSNDVRDAALYSSORRHETESAGRGYPVEGAGSGAIRG 240  
Db 181 MKNGYSTENREWTNTLGFVGSNDVRDAALYSSORRHETESAGRGYPVEGAGSGANIRG 240  
Qy 241 SSRGIPDPSKHXYNFKIAYQINDKRIKIPSPNGQGHNYTTEESVNLTFASWREADD 300  
Db 241 SARGIPDPSKHXYNFKIAYQINDKRIKIPSPNGQGHNYTTEESVNLTFASWREADD 300  
Qy 301 VNRERNANLFYEWTPDNLSSLKADPDYOTTKVAANVNGSGFP-TDYSTWTRVNNOKDL 359  
Db 301 VNRERNANLFYEWTPDNLSSLKADPDYOTTKVAANVNGSGFP-TDYSTWTRVNNOKDL 360  
Qy 360 ENIYNRSMDFRFRFTLRMDSPQLQL-CGQHRLSLKTFFASRRREFENLRDYYFSEVSR 418  
Db 361 DEIYNRSMDFRFRFTLRMDSPQLQL-CGQHRLSLKTFFASRRREFENLRDYYFSEVSR 420  
Qy 419 TTSSIQHPVKTNTNYGFSLSQIQWNVFSSRADIRYDHTKMTPOELNAECHACDKTPPAA 478  
Db 421 TTSSIQHPVKTNTNYGFSLSQIQWNVFSSRADIRYDHTKMTPOELNAECHACDKTPPAA 480  
Qy 479 NTYKMGSGFVGLAAQLNAQWVGYDITSGYRVPNASEVYFTYHSGNWLNPNLKABRS 538  
Db 481 NTYKMGSGFVGLAAQLNAQWVGYDITSGYRVPNASEVYFTYHSGNWLNPNLKABRS 540  
Qy 539 TTHTLSLQGRSEKGTLDANLYQNNYRNFLESEOKLTTSGDVCTQMNYYGMCNPNYSEK 598  
Db 541 TTHTLSLQGRSEKGTLDANLYQNNYRNFLESEOKLTTSGDVCTQMNYYGMCNPNYSEK 600  
Qy 599 PEMQMONIDKARIRGLTGLRLNVTKVASFVPEGWKLFGLSGLYAKSLSGDNLSTQPP 658  
Db 601 LEWQMONIDKARIRGLTGLRLNVTKVASFVPEGWKLFGLSGLYAKSLSGDNLSTQPP 660  
Qy 659 KVIAGVDYSPSEKMGVFSRLTYLGAKKADAQYTVYENKRGCTPLQKKVKDYPWLNKSA 718  
Db 661 KVIAGIDYSPSEKMGVFSRLTYLGAKKADAQYTVYENKRGCTPLQKKVKDYPWLNKSA 720  
Qy 719 YVFDMYGFYKLAQNLTLRAGVYVFNKRYTTWDSLRGLYSYSTTNAVDRDGKGLDRYAS 778  
Db 721 YVFDMYGFYKLAQNLTLRAGVYVFNKRYTTWDSLRGLYSYSTTNAVDRDGKGLDRYAS 780  
Qy 779 GRNYAVSLDWKF 790  
Db 781 SRNYAVSLDWKF 792  
RESULT 4  
US-08-817-707-2  
; Sequence 2, Application US/08817707  
; Patent No. 6277382  
; GENERAL INFORMATION:  
; APPLICANT: Stojiljkovic, Igor  
; APPLICANT: So, Magdalene  
; APPLICANT: Hwa, Vivian  
; APPLICANT: Heffron, Fred  
; APPLICANT: Nassif, Xavier  
; TITLE OF INVENTION: No. 6277382el Bacterial Hemoglobin Receptor  
; TITLE OF INVENTION: Genes and Uses  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive, 32nd Floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/817,707



Db 1 MKPLHMLPIAALVGSIFGNPVLAADEAAETTPVKAIEKEVRVKDQLNAPATVERVNLGR 60  
QY 61 IQEEMIRNDKDLVRYSTDVGLSDSGRHQKGFVAVRGVEGNRVGVSIDGVSILPSEENSLYA 120  
Db 61 IQEEMIRNDKDLVRYSTDVGLSDSGRHQKGFVAVRGVEGNRVGVSIDGVSILPSEENSLYA 120  
QY 121 RYGNFNSSRLSDPELVNRIEIAKAGDSFNTGSGALGGGVNVTQLQGHDLDDDRQFGVM 180  
Db 121 RYGNFNSSRLSDPELVNRIEIAKAGDSFNTGSGALGGGVNVTQLQGHDLDDDRQFGVM 180  
QY 181 MKNGYSRRNREWTNTLFGVSNDRVDAALYSORRGHETESAGERGYPVEGAGSGAIRG 240  
Db 181 MKNGYSRRNREWTNTLFGVSNDRVDAALYSORRGHETESAGERGYPVEGAGSGAIRG 240  
QY 241 SRGIPDPKSKHYHFLGKIAYQINDKHRIKGFSPNGQOQHNTYIESYNLTAASSWREADD 300  
Db 241 SRGIPDPKSKHYHFLGKIAYQINDKHRIKGFSPNGQOQHNTYIESYNLTAASSWREADD 300  
QY 301 VNRRNANLFYEWTPDSNWLSSLKADFDYQTTKVAAVNNKGSFPPTDYSTWTRYNKOLE 360  
Db 301 VNRRNANLFYEWTPDSNWLSSLKADFDYQTTKVAAVNNKGSFPPTDYSTWTRYNKOLE 360  
QY 361 NIYNRSMDFRFRFTLRMDSQPLQGGQHRLSLKTFASSRREPENLRDDYYSERSVRTT 420  
Db 361 NIYNRSMDFRFRFTLRMDSQPLQGGQHRLSLKTFASSRREPENLRDDYYSERSVRTT 420  
QY 421 SSIQHPVKTTNYGFSLSDOIQWNVFSSRADIRYDHTKMTPOELNAECHACDKTPPAANT 480  
Db 421 SSIQHPVKTTNYGFSLSDOIQWNVFSSRADIRYDHTKMTPOELNAECHACDKTPPAANT 480  
QY 481 YKGSFGFVGLAQLNQAQWVGYDITSGYRVPNASEVYFTYNNHSGGNWLPNPKLKAERSTT 540  
Db 481 YKGSFGFVGLAQLNQAQWVGYDITSGYRVPNASEVYFTYNNHSGGNWLPNPKLKAERSTT 540  
QY 541 HTLSLQGRSEKGTLDANLYQNNYRNFPLSBEQKLTTSBGVGTQMMYYGMCNPNYSEKPE 600  
Db 541 HTLSLQGRSEKGTLDANLYQNNYRNFPLSBEQKLTTSBGVGTQMMYYGMCNPNYSEKPE 600  
QY 601 WQMONIDKARIRGLELTGRLNVTKVASFVPEGWKLFGLSLGYAKSKLSDGNSLLSTQPPKV 660  
Db 601 WQMONIDKARIRGLELTGRLNVTKVASFVPEGWKLFGLSLGYAKSKLSDGNSLLSTQPPKV 660

RESULT 2

US-08-537-361E-8  
; Sequence 8, Application US/08537361E  
; Patent No. 6121037  
; GENERAL INFORMATION:  
; APPLICANT: Stojiljkovic, Igor  
; APPLICANT: So, Magdalene  
; APPLICANT: Hwa, Vivian  
; APPLICANT: Heffron, Fred  
; APPLICANT: Nassif, Xavier  
; TITLE OF INVENTION: No. 6121037el Bacterial Hemoglobin Receptor  
; TITLE OF INVENTION: Genes and Uses  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive, 32nd Floor  
; CITY: Chicago  
; STATE: Illinois

COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/537,361E  
FILING DATE: 02-OCT-1995  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6121037nan, Kevin E  
REGISTRATION NUMBER: 35,303  
REFERENCE/DOCKET NUMBER: 94,784-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-913-0001  
TELEFAX: 312-913-0002  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 790 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-537-361E-8  
Query Match 99.3%; Score 4171; DB 3; Length 790;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 784; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MKPLHMLPIAALVGSIFGNPVLAADEAAETTPVKAIEKEVRVKDQLNAPATVERVNLGR 60  
Db 1 MKPLHMLPIAALVGSIFGNPVLAADEAAETTPVKAIEKEVRVKDQLNAPATVERVNLGR 60  
QY 61 IQEEMIRNDKDLVRYSTDVGLSDSGRHQKGFVAVRGVEGNRVGVSIDGVSILPSEENSLYA 120  
Db 61 IQEEMIRNDKDLVRYSTDVGLSDSGRHQKGFVAVRGVEGNRVGVSIDGVSILPSEENSLYA 120  
QY 121 RYGNFNSSRLSDPELVNRIEIAKAGDSFNTGSGALGGGVNVTQLQGHDLDDDRQFGVM 180  
Db 121 RYGNFNSSRLSDPELVNRIEIAKAGDSFNTGSGALGGGVNVTQLQGHDLDDDRQFGVM 180  
QY 181 MKNGYSRRNREWTNTLFGVSNDRVDAALYSORRGHETESAGERGYPVEGAGSGAIRG 240  
Db 181 MKNGYSRRNREWTNTLFGVSNDRVDAALYSORRGHETESAGERGYPVEGAGSGAIRG 240  
QY 241 SRGIPDPKSKHYHFLGKIAYQINDKHRIKGFSPNGQOQHNTYIESYNLTAASSWREADD 300  
Db 241 SRGIPDPKSKHYHFLGKIAYQINDKHRIKGFSPNGQOQHNTYIESYNLTAASSWREADD 300  
QY 301 VNRRNANLFYEWTPDSNWLSSLKADFDYQTTKVAAVNNKGSFPPTDYSTWTRYNKOLE 360  
Db 301 VNRRNANLFYEWTPDSNWLSSLKADFDYQTTKVAAVNNKGSFPPTDYSTWTRYNKOLE 360  
QY 361 NIYNRSMDFRFRFTLRMDSQPLQGGQHRLSLKTFASSRREPENLRDDYYSERSVRTT 420  
Db 361 NIYNRSMDFRFRFTLRMDSQPLQGGQHRLSLKTFASSRREPENLRDDYYSERSVRTT 420  
QY 421 SSIQHPVKTTNYGFSLSDOIQWNVFSSRADIRYDHTKMTPOELNAECHACDKTPPAANT 480  
Db 421 SSIQHPVKTTNYGFSLSDOIQWNVFSSRADIRYDHTKMTPOELNAECHACDKTPPAANT 480  
QY 481 YKGSFGFVGLAQLNQAQWVGYDITSGYRVPNASEVYFTYNNHSGGNWLPNPKLKAERSTT 540  
Db 481 YKGSFGFVGLAQLNQAQWVGYDITSGYRVPNASEVYFTYNNHSGGNWLPNPKLKAERSTT 540  
QY 541 HTLSLQGRSEKGTLDANLYQNNYRNFPLSBEQKLTTSBGVGTQMMYYGMCNPNYSEKPE 600  
Db 541 HTLSLQGRSEKGTLDANLYQNNYRNFPLSBEQKLTTSBGVGTQMMYYGMCNPNYSEKPE 600  
QY 601 WQMONIDKARIRGLELTGRLNVTKVASFVPEGWKLFGLSLGYAKSKLSDGNSLLSTQPPKV 660  
Db 601 WQMONIDKARIRGLELTGRLNVTKVASFVPEGWKLFGLSLGYAKSKLSDGNSLLSTQPPKV 660

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: May 6, 2005, 14:49:54 ; Search time 43 Seconds  
(without alignments)  
1371.460 Million cell updates/sec

Title: US-09-665-358-8

Perfect score: 4200

Sequence: 1 MKPLHMLPIAALVGSIFGNP.....GLDRYASGRNVAVSLDWKF 790

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pdp.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pdp.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pdp.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pdp.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pdp.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description          |
|------------|--------|-------------|--------|----|----------------------|
| 1          | 4187   | 99.7        | 790    | 3  | US-08-817-707-8      |
| 2          | 4171   | 99.3        | 790    | 3  | US-08-537-361E-8     |
| 3          | 3814   | 90.8        | 792    | 3  | US-08-990-470A-2     |
| 4          | 3814   | 90.8        | 792    | 3  | US-08-817-707-2      |
| 5          | 3812   | 90.8        | 792    | 1  | US-08-326-670A-2     |
| 6          | 3810   | 90.7        | 790    | 3  | US-08-817-707-6      |
| 7          | 3806   | 90.6        | 792    | 3  | US-08-537-361E-6     |
| 8          | 3799   | 90.5        | 792    | 3  | US-08-537-361E-2     |
| 9          | 3668.5 | 87.3        | 791    | 3  | US-08-817-707-4      |
| 10         | 3665.5 | 87.3        | 791    | 3  | US-08-537-361E-4     |
| 11         | 753.5  | 17.9        | 967    | 4  | US-09-809-665A-32    |
| 12         | 737.5  | 17.6        | 669    | 4  | US-09-809-665A-105   |
| 13         | 702.5  | 16.7        | 789    | 4  | US-09-252-991A-27011 |
| 14         | 666    | 15.9        | 944    | 2  | US-08-867-941-24     |
| 15         | 666    | 15.9        | 944    | 3  | US-09-074-658-24     |
| 16         | 651.5  | 15.5        | 944    | 2  | US-08-867-941-23     |
| 17         | 651.5  | 15.5        | 944    | 3  | US-09-074-658-23     |
| 18         | 642.5  | 15.3        | 971    | 3  | US-03-405-728-2      |
| 19         | 641    | 15.3        | 941    | 3  | US-09-074-658-75     |
| 20         | 640    | 15.2        | 931    | 3  | US-08-624-655A-2     |
| 21         | 631    | 15.0        | 1052   | 4  | US-08-778-570B-14    |
| 22         | 631    | 15.0        | 1052   | 4  | US-09-059-584-14     |
| 23         | 631    | 15.0        | 1070   | 3  | US-08-613-009A-11    |
| 24         | 631    | 15.0        | 1070   | 4  | US-08-778-570B-13    |
| 25         | 631    | 15.0        | 1070   | 4  | US-09-059-584-13     |
| 26         | 629.5  | 15.0        | 1053   | 3  | US-08-613-009A-8     |
| 27         | 629.5  | 15.0        | 1053   | 4  | US-08-778-570B-10    |

|    |       |      |      |   |                    |                   |
|----|-------|------|------|---|--------------------|-------------------|
| 28 | 629.5 | 15.0 | 1053 | 4 | US-09-059-584-10   | Sequence 10, Appl |
| 29 | 629.5 | 15.0 | 1074 | 3 | US-08-613-009A-7   | Sequence 7, Appli |
| 30 | 629.5 | 15.0 | 1074 | 4 | US-08-778-570B-9   | Sequence 9, Appli |
| 31 | 629.5 | 15.0 | 1074 | 4 | US-09-059-584-9    | Sequence 9, Appli |
| 32 | 625.5 | 14.9 | 1076 | 2 | US-08-867-941-19   | Sequence 19, Appl |
| 33 | 625.5 | 14.9 | 1076 | 3 | US-09-074-658-19   | Sequence 19, Appl |
| 34 | 606   | 14.4 | 930  | 4 | US-08-753-750B-2   | Sequence 2, Appli |
| 35 | 597.5 | 14.2 | 911  | 1 | US-08-487-890A-107 | Sequence 107, App |
| 36 | 597.5 | 14.2 | 911  | 2 | US-08-478-435-107  | Sequence 107, App |
| 37 | 597.5 | 14.2 | 911  | 2 | US-08-337-483-107  | Sequence 107, App |
| 38 | 597.5 | 14.2 | 911  | 2 | US-08-478-373-107  | Sequence 107, App |
| 39 | 597.5 | 14.2 | 911  | 3 | US-08-474-671-107  | Sequence 107, App |
| 40 | 597.5 | 14.2 | 911  | 3 | US-08-483-577A-107 | Sequence 107, App |
| 41 | 597.5 | 14.2 | 911  | 3 | US-08-897-438-107  | Sequence 107, App |
| 42 | 597.5 | 14.2 | 911  | 3 | US-08-637-654-107  | Sequence 107, App |
| 43 | 597.5 | 14.2 | 911  | 3 | US-08-649-518-107  | Sequence 107, App |
| 44 | 590   | 14.0 | 914  | 1 | US-08-487-890A-11  | Sequence 11, Appl |
| 45 | 590   | 14.0 | 914  | 2 | US-08-478-435-11   | Sequence 11, Appl |

## ALIGNMENTS

RESULT 1  
US-08-817-707-8  
; Sequence 8, Application US/08817707  
; Patent No. 6277382  
; GENERAL INFORMATION:  
; APPLICANT: Stojiljkovic, Igor  
; APPLICANT: So, Magdalene  
; APPLICANT: Hwa, Vivian  
; APPLICANT: Heffron, Fred  
; APPLICANT: Nassif, Xavier  
; TITLE OF INVENTION: No. 6277382el Bacterial Hemoglobin Receptor  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff  
; STREET: 300 South Wacker Drive, 32nd Floor  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/817,707  
; FILING DATE: 19-AUG-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 6277382nan, Kevin E  
; REGISTRATION NUMBER: 35,303  
; REFERENCE/DOCKET NUMBER: 94,784-J  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-913-0001  
; TELEFAX: 312-913-0002  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 790 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-817-707-8

Query Match 99.7% Score 4187; DB 3; Length 790;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 786; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKPLHMLPIAALVGSIFGNPVLAADEAATTTPKAEIKVRVKDQNPATVERVNLGR 60  
|||||

| Query Match           | 18.1%            | Score 760  | DB 8       | Length 1046 |
|-----------------------|------------------|--|------------|-------------|
| Best Local Similarity | 23.8%            | Pred. No. 6.9e-54  |            |             |
| Matches 248           | Conservative 141 | Mismatches 317   | Indels 338 | Gaps 30     |
| Qy                    | 45               | DQLNAPATVERNL-----CRIOQMIRDNKDLVRYSTDVGLSDSGR-HQKGF            | 91         |             |
| Db                    | 43               | EQINVSGSTTINVKEKKVGETQISAKKLAKQASDSDLVRVETGTVVETGRTGASGY       | 102        |             |
| Qy                    | 92               | AVRGEVGNRVGVSIDGV-----SLPDSSENSLYARYGNFNSRSLSIDPELVNRIEIAKGD   | 147        |             |
| Db                    | 103              | AVRGVDENRVGIMVDGLRQAETLSSQGFELFEGYGNFNTRNSIEIENVKTATITKGAD     | 162        |             |
| Qy                    | 148              | SFNTGSGALGGVNYQTLQGHDLDDRRQGVGMWKNQYSRNRBWTWTLFGVSNDRVDA       | 207        |             |
| Db                    | 163              | SLKSGSGALGGSVIFETKDARDYLI-DKDYLSYKRGYQTMNQNLKTLTLAGRSKKFDI     | 221        |             |
| Qy                    | 208              | ALLYSQRSHETESAGRGYPVEGAGSGAIRGSRGIPDPKSHKHNFGLKLAYQINDK        | 267        |             |
| Db                    | 222              | LVDTRKDGHEIENYDKIIPNQADLSAV--GPTREKADPYQITROSTLIKGFQPNEN       | 279        |             |
| Qy                    | 268              | HRIGP-----SFNGQQGHNYTIESYNLTASSWREADDVNRNRNANLFYE-             | 312        |             |
| Db                    | 280              | HRLSVALDDSTLETKGMDLSYIFNQCKTCEEKY-----GERVINDQSRKNIQPSYEN      | 334        |             |
| Qy                    | 313              | -----WTPDSNMLSSLKA-----DFDYQTTKVAANNNKGSF-----PTDYS            | 348        |             |
| Db                    | 335              | FSQTPFDWHIKLSYSSQKIITNKARSDEYCHQSTCAGVRNPQGLHLVBEGGIYKIKQDNQ   | 394        |             |
| Qy                    | 349              | TWRTNYNQ-----KOLEN-----  | 361        |             |
| Db                    | 395              | EFTGKYNIGLELQWKGEDVSNVDVTGGSLDSVLINCEKLNCKNKKFQVFGVDENNDK      | 454        |             |
| Qy                    | 362              | -----  | 369        |             |
| Db                    | 455              | YNYEDRNITIKELNGKKYGEISLKSASGNYGLTQYESAKFLPKSHGYSYTDVFNDRDLNT   | 514        |             |
| Qy                    | 370              | RFRKFTLRMSQ-----PLQGGQHRLSLKTFASSRREFENLRDDY---YPSERVSR        | 418        |             |
| Db                    | 515              | HTQCIKLDDKEFHLWHTQHLKYGGLYEKTLSKSMVNHQYVTVANVQWAGNFFCNKFPD     | 574        |             |
| Qy                    | 419              | TTSSIOH-----PVKTTNYGFSLSDOIQWMDVFSRRADIR                       | 453        |             |
| Db                    | 575              | GTYPVEHTPADHSAYKCSLMSNKGKDTYLIPTTKNNVLYFGDNVQLTSLWGLDLNLR      | 634        |             |
| Qy                    | 454              | YDHTKMTPOBELNAECHACDKTPPAAN-----                               | 481        |             |
| Db                    | 635              | YDHVKYLP-----SYDKNI PVPNGLI TGLFKFKFSKEYVYGSKYISIPGYENCITYDT   | 686        |             |
| Qy                    | 482              | -----KGWSGFGVLAALQNAWHGYDI-----TSGYRVNAGSEVYFTYN               | 521        |             |
| Db                    | 687              | PCYKKNPEDNLALLRRKTDYKHSHYNLGLNLDPTNLWRVOLKYANGFRAPTSDEIYMTFK   | 746        |             |
| Qy                    | 522              | HGSGNWLPNPLKAERSTHTLSLQGRSEKGTLDANLYQNNRYNRLFSEOKLTTSGDVG      | 581        |             |
| Db                    | 747              | HPQFSIQNTDLKAETSKTEVAFTFYKNSYIITLNAFONDYRNFID-----LVEVGE-      | 799        |             |
| Qy                    | 582              | TQMYNYGCMSPYSE-----KPEWOMONTDKARIRGELTGRLVNTKVASFVPEGKLF       | 636        |             |
| Db                    | 800              | -----RPIEESAIKYPPHQNQRDRVRGIEIASRLMEGDLFEKL-QGFHLG             | 847        |             |
| Qy                    | 637              | GSLGYAKSKLSGDNSL-----  | 668        |             |
| Db                    | 848              | YKFTYQGRIR-DNGLHPKYKEFLELNKDEHPEVEIARQPQMNALQPTTSVNYGYDA       | 906        |             |
| Qy                    | 669              | PSEKWGVFSRLTYLGAKAKDA-----QYT-----YVENKGRGTPLOKKVK             | 709        |             |
| Db                    | 907              | PSQKWGVDMYITNVAACKAKDSFNSQWTSMAVAREKEQVTONATDIPATKANG-----KDKV | 962        |             |
| Qy                    | 710              | DYP--WLKSAVYFDMYGFYKLANLTLRAGVYVFNRYKTTWDSLRLGYSYSTTNNAVDR     | 767        |             |
| Db                    | 963              | DSRGLWRNRYTVTIDTAYKPIKLNLTFTAGVYNTNRKLYLTWDSAKSVRHLGTINRVKT    | 1022       |             |
| Qy                    | 768              | D-GKGLDRYRASGRNYAVSLDWKF                                       | 790        |             |

|    |   |   |                     |           |             |  |  |  |  |  |
|----|---|---|---------------------|-----------|-------------|--|--|--|--|--|
| CC | comprising the protein, nucleic acid or antibody is useful for the    |   |                     |           |             |  |  |  |  |  |
| CC | manufacture of a medicament for treating or preventing N. gonorrhoeae |   |                     |           |             |  |  |  |  |  |
| CC | infection, this may be in the form of a vaccine or gene therapy.      |   |                     |           |             |  |  |  |  |  |
| CC | Sequences given in records ABP76736-ABP81046 represent nucleic acid   |   |                     |           |             |  |  |  |  |  |
| CC | molecules of the invention  |   |                     |           |             |  |  |  |  |  |
| XX |   |   |                     |           |             |  |  |  |  |  |
| SQ | Sequence 288 AA;  |   |                     |           |             |  |  |  |  |  |
|    | Query Match   | 34.8%;  | Score 1462.5;       | DB 6;     | Length 288; |  |  |  |  |  |
|    | Best Local Similarity   | 95.5%;  | Pred. No. 1.2e-113; |           |             |  |  |  |  |  |
|    | Matches 274;  | Conservative 5;   | Mismatches 7;       | Indels 1; | Gaps 1;     |  |  |  |  |  |
| QY | 391   | LSLKTFSARREFENLNDDYYFSERSVGRSTSSIOHPVKTTNYGSLSDQIQWNVDFSRA    | 450                 |           |             |  |  |  |  |  |
| DB | 1   | LSFPTFSARREFENLNDDYYFSQIQRTSSIOHPVKTTNYGSLSDQIQWNVDFSRA       | 60                  |           |             |  |  |  |  |  |
| QY | 451   | DIRYDHTKMTPOELNAEACHADCKTPPAAANTYKGWSGFVGLAAQLNQAHVGYDITSGYRV | 510                 |           |             |  |  |  |  |  |
| DB | 61  | DIRYDHTKMTPOELNAEACHADCKTPPAAANTYKGWSGFVGLAAQLNQAHVGYDITSGYRV | 120                 |           |             |  |  |  |  |  |
| QY | 511   | PNASEVYFTYNHGSGNWLPHNPKAERSTHTTSLQGRSEKGTLDANLYQNNYRNFISEE    | 570                 |           |             |  |  |  |  |  |
| DB | 121   | PNASEVYFTYNHGSGNWLPHNPKAERSTHTTSLQGRSEKGTLDANLYQNNYRNFISEE    | 180                 |           |             |  |  |  |  |  |
| QY | 571   | QKLTTSGDVGCTQWNYYYGHCNSPNYSKEPQWQMONIDKARIRGLELTGRLNVTKVASFVP | 630                 |           |             |  |  |  |  |  |
| DB | 181   | QKLTTSGDVGCTQWNYYYGHCNSPNYSKEPQWQMONIDKARIRGLELTGRLNVTKVASFVP | 240                 |           |             |  |  |  |  |  |
| QY | 631   | EGWKLFGSLGYAKSKLSGDNSLLSTQPPKVIAGVDYSPSEKWKGVFS               | 677                 |           |             |  |  |  |  |  |
| DB | 241   | EGWKLFGSLGYAKSKLSGDNSLLSTQPPKVIAGVDYSEARA-KNGVCS              | 286                 |           |             |  |  |  |  |  |

|          |   |
|----------|---|
| RESULT 6 |   |
| ABP77003 |   |
| ID       | ABP77003 standard; protein; 284 AA.                                     |
| XX       | AC  |
| XX       | ABP77003;   |
| XX       | AC  |
| XX       | 07-MAR-2003 (first entry)   |
| DT       |   |
| XX       |   |
| DE       | N. gonorrhoeae amino acid sequence SEQ ID 536.                          |
| XX       |   |
| KW       | Antibacterial; infection; vaccine; gene therapy.                        |
| XX       |   |
| OS       | Neisseria gonorrhoeae.  |
| XX       |   |
| PN       | WO200279243-A2.   |
| XX       |   |
| PD       | 10-OCT-2002.  |
| XX       |   |
| PF       | 12-FEB-2002; 2002WO-IB002069.   |
| XX       |   |
| PR       | 12-FEB-2001; 2001GB-00003424.   |
| XX       |   |
| PA       | (CHIR-) CHIRON SPA.   |
| XX       |   |
| PI       | Fontana MR, Pizza M, Masignani V, Monaci E;                             |
| XX       |   |
| DR       | WPI; 2003-058415/05.  |
| DR       | N-PSDB; ABZ37973.   |
| XX       |   |
| PT       | New protein from Neisseria gonorrhoeae, useful for the manufacture of a |
| PT       | medicament for treating or preventing N. gonorrhoeae infection.         |
| PT       |   |

|          |   |  |
|----------|---|--|
| CC       | Sequences given in records ABP76736-ABP81046 represent nucleic acid       |  |
| CC       | molecules of the invention  |  |
| XX       |   |  |
| SQ       | Sequence 284 AA;  |  |
|          | Query Match 34.4%; Score 1445; DB 6; Length 284;                          |  |
|          | Best Local Similarity 98.6%; Pred. No. 3.5e-112;                          |  |
|          | Matches 280; Conservative 1; Mismatches 3; Indels 0; Gaps 0;              |  |
| Qy       | 1 MKPEHMLPTAALVGSIFGNPVLAADEAAATETTPVKAEIKEVRVKQQLNAPATVERVNLGR 60        |  |
| Db       | 1 MKPEHMLPTAALVGSIFGNPVLAADEAAATETTPVKAEIKEVRVKQQLNAPATVERVNLDR 60        |  |
| Qy       | 61 IQQEMIRDNDKLVRYSTDVGLSDSGRHKGFVARGVGEGRNVGVSIDGVSLPDSEENSLYA 120       |  |
| Db       | 61 IQQEMIRDNDKLVRYSTDVGLSDSGRHKGFVARGVGEGRNVGVSIDGVSLPDSEENSLYA 120       |  |
| Qy       | 121 RYGNFNSSRLSIDPELVRNIETAKGADSFNTGSGALGGGVNYQTLOGHDLILLDDRQFGVM 180     |  |
| Db       | 121 RYGNFNSSRLSIDPELVRNIETAKGADSFNTGSGALGGGVNYQTLOGHDLILLDDRQFGVM 180     |  |
| Qy       | 181 MKNGYSRRNEWNTLFGFVGSNDNRDVAALLYSORRGHETESAGERGYPVEGAGSGAIIRG 240      |  |
| Db       | 181 MKNGYSRRNEWNTLFGFVGSNDNRDVAALLYSORRGHETESAGERGYPVEGAGSGANIRG 240      |  |
| Qy       | 241 SSRGIPDPKSKHYHFLGKIAYOINDKHRIGPSFNGQQGHNYTI 284                       |  |
| Db       | 241 SARGIPDPKSKHYHFLGKIAYOINDKHRIGASFNQQGHNYTI 284                        |  |
| RESULT 7 |   |  |
| ADT05707 | ID ADT05707 standard; protein; 1046 AA.                                   |  |
| XX       | ADT05707;   |  |
| AC       |   |  |
| XX       |   |  |
| DT       | 02-DEC-2004 (first entry)   |  |
| XX       |   |  |
| DE       | Haemophilus influenzae (NTHI) protein - SEQ ID 743.                       |  |
| XX       |   |  |
| KW       | middle ear bacterial infection; nasopharynx bacterial infection.          |  |
| XX       |   |  |
| OS       | Haemophilus influenzae.   |  |
| XX       |   |  |
| PN       | WO2004078949-A2.  |  |
| XX       |   |  |
| PD       | 16-SEP-2004.  |  |
| XX       |   |  |
| PF       | 05-MAR-2004; 2004WO-US007001.   |  |
| XX       |   |  |
| PR       | 06-MAR-2003; 2003US-0453134P.   |  |
| XX       |   |  |
| PA       | (CHIL-) CHILDRENS HOSPITAL INC.   |  |
| XX       |   |  |
| PI       | Bakaletz LO, Munson RS, Dyer DW;  |  |
| XX       |   |  |
| DR       | WPI; 2004-662422/64.  |  |
| DR       | N-PSDB; ADT05706.   |  |
| XX       |   |  |
| PT       | New polynucleotides of nontypeable strain of Haemophilus influenzae,      |  |
| PT       | useful for treating or preventing NTHi bacterial infections of the middle |  |
| PT       | ear and/or nasopharynx.   |  |
| XX       |   |  |
| PS       | Claim 3; SEQ ID NO 743; 88pp; English.                                    |  |
| XX       |   |  |
| CC       | The invention comprises nucleotide sequences (genes) from the genome of a |  |
| CC       | nontypeable strain of Haemophilus influenzae (NTHI). The NTHi DNA         |  |
| CC       | sequences of the invention are useful for treating or preventing NTHi     |  |
| CC       | bacterial infections of the middle ear and/or nasopharynx. The present    |  |
| CC       | amino acid sequence represents an NTHi protein of the invention.          |  |
| XX       |   |  |
| SQ       | Sequence 1046 AA;   |  |

XX The invention comprises nucleotide sequences (genes) from the genome of a  
CC nontypeable strain of *Haemophilus influenzae* (NTHI<sup>1</sup>). The NTHI DNA  
CC sequences of the invention are useful for treating or preventing Nthi  
CC bacterial infections of the middle ear and/or nasopharynx. The present  
CC amino acid sequence represents an NTHi protein of the invention.  
XX  
SQ Sequence 1046 AA;  
SQ



QY 420 TSSIQHPVKTNTYGFSLSDQIQWNVDFSSRADIRYDHTKMTPOELNAEACHADCTPPAAN 479  
 Db 420 TSSIQHPVKTNTYGFSLSDQIQWNVDFSSRADIRYDHTKMTPOELNAEACHADCTPPAAN 479  
 QY 480 TYKWSGFGVGLAAQLNQAHVGYDITSGYRVPNASEVFTYNNHSGNWLPPNPKAERST 539  
 Db 480 TYKWSGFGVGLAAQLNQAHVGYDITSGYRVPNASEVFTYNNHSGNWLPPNPKAERST 539  
 QY 540 THTLSLQGRSEKGTLDANLYQNNYENFLSEBQKLTTSQDVCTQMNYYGWCSPYSEKP 599  
 Db 540 THTLSLQGRSEKGTLDANLYQNNYENFLSEBQKLTTSQDVCTQMNYYGWCSPYSEKP 599  
 QY 600 EQWQONIDKARIRGIELTGRNLVTVKASFPVEGKLFGLSLGYAKSKLSDGNSLLSTQPPK 659  
 Db 600 EQWQONIDKARIRGIELTGRNLVTVKASFPVEGKLFGLSLGYAKSKLSDGNSLLSTQPPK 659  
 QY 660 VIAGVDYSPSEKMGVFSRLTYLGAKKAKDAQYTYVYENKGRGTPLQKKVDPYMLNKSAY 719  
 Db 660 VIAGVDYSPSEKMGVFSRLTYLGAKKAKDAQYTYVYENKGRGTPLQKKVDPYMLNKSAY 719  
 QY 720 VEDMGVGFYKAKNLTLAGVNVNENKYYTWDLSRLGYSYSTTNAVDRDGLDRYRAG 779  
 Db 720 VEDMGVGFYKAKNLTLAGVNVNENKYYTWDLSRLGYSYSTTNAVDRDGLDRYRAG 779  
 QY 780 RNYAVSLDWKF 790  
 Db 780 RNYAVSLEWKF 790

## RESULT 3

AAR95567  
 ID AAR95567 standard; protein; 792 AA.

XX AAR95567;

DT 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 28-AUG-1996 (first entry)

XX N. meningitidis serotype B haemoglobin receptor.

XX Bacterial haemoglobin receptor; hmbR gene; antibody; vaccine; meningitis.

XX Neisseria meningitidis serogroup B.

XX W09612020-A2.

XX 25-APR-1996.

XX 17-OCT-1995; 95WO-US013623.

XX 18-OCT-1994; 94US-00326670.

XX 02-OCT-1995; 95US-00537361.

XX (UYOR-) UNIV OREGON HEALTH SCI.

XX Stojiljkovic I, So M, Hwa V, Heffron F, Nassif X;

XX WPI; 1996-222006/22.

XX N-PSDB; AAT26999.

XX DNA encoding Neisseria haemoglobin receptor proteins - for use in preparing polypeptide(s) and antibodies for vaccines for e.g. meningitis.

XX Claim 4; Page 58-61; 104pp; English.

XX A bacterial haemoglobin receptor (AAR95567) was identified as the product of the hmbR gene (AAT26999) of N. meningitidis serotype B. The haemoglobin receptor is thought to be involved in hemin utilisation by the bacterium. This dependence on host iron stores is a potential route for therapeutic intervention strategies. The receptor can be expressed in transformed hosts, e.g. attenuated Salmonella cells. It is useful as a

CC vaccine for meningitis, or can be used to raise antibodies of diagnostic appln. Similar receptor proteins were obtd. from N. meningitidis CC serotypes A (AAR95566) and C (AAR95565) and from N. gonorrhoeae CC (AAR95568). (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to correct OS field.)

XX SQ Sequence 792 AA;

Query Match 90.8%; Score 3812; DB 2; Length 792;

Best Local Similarity 91.2%; Pred. No. 8e-310;

Matches 722; Conservative 25; Mismatches 43; Indels 2; Gaps 2;

QY 1 MKPLHMLPIAALVGSIFGNPVLAADEAATETTPVKAIEKEVRVKDQLNAPATVERVNLGR 60  
 Db 1 MKPLHMLPIAALVGSIFGNPVLAADEAATETTPVKAIEKEVRVKDQLNAPATVERVNLGR 60  
 QY 61 IQQEMIRNDKOLVRYSTDVGLSDSGRHQKQFAVRGVEGNRVGVSIDGVSLEPSENSLYA 120  
 Db 61 IQQEMIRNDKOLVRYSTDVGLSDSGRHQKQFAVRGVEGNRVGVSIDGVSLEPSENSLYA 120  
 QY 121 RYGFNSRLSIDPELVNIEIAKADSFNTGSGALGGGVNYOTLQGHDLILLDDRFQGM 180  
 Db 121 RYGFNSRLSIDPELVNIEIAKADSFNTGSGALGGGVNYOTLQGHDLILLDDRFQGM 180  
 QY 181 MKNGYSSRNREWTNTLGFVSNDRVDAALLYSORRGHETESAGERYGVEGAGSGAIIRG 240  
 Db 181 MKNGYSSRNREWTNTLGFVSNDRVDAALLYSORRGHETESAGERYGVEGAGSGAIIRG 240  
 QY 241 SSRGIPDPSPKHYNFLGKIAYQINDKRIHIGPSFNGQOQHNYTIEESYNLTASSWREADD 300  
 Db 241 SARGIPDPSPKHYSFLGKIAYQINDKRIHIGPSFNGQOQHNYTIEESYNLTASSWREADD 300  
 QY 301 VNRERNANLFYEWTPDSNWLSSLKADFDYQTKVAANVKNKSGFP-TDYSTWTRVYNQKDL 359  
 Db 301 VNRERNANLFYEWTPDSNWLSSLKADFDYQTKVAANVKNKSGFP-TDYSTWTRVYNQKDL 359  
 QY 360 ENIYNRSMDTRFKFTLRMDSDQLQ-LGGHRLSLKTFASRRFENLNRDDYFYSERVSR 418  
 Db 360 DEIYNRSMDTRFKFTLRMDSDQLQ-LGGHRLSLKTFASRRFENLNRDDYFYSERVSR 418  
 QY 419 TTSSIQHPVKTNTYGFSLSDQIQWNVDFSSRADIRYDHTKMTPOELNAEACHADCTPPAA 478  
 Db 419 TTSSIQHPVKTNTYGFSLSDQIQWNVDFSSRADIRYDHTKMTPOELNAEACHADCTPPAA 478  
 QY 479 NTKGWSGFGVGLAAQLNQAHVGYDITSGYRVPNASEVFTYNNHSGNWLPPNPKAERS 538  
 Db 479 NTKGWSGFGVGLAAQLNQAHVGYDITSGYRVPNASEVFTYNNHSGNWLPPNPKAERS 538  
 QY 481 NTKGWSGFGVGLAAQLNQAHVGYDITSGYRVPNASEVFTYNNHSGNWLPPNPKAERT 540  
 Db 481 NTKGWSGFGVGLAAQLNQAHVGYDITSGYRVPNASEVFTYNNHSGNWLPPNPKAERT 540  
 QY 539 THTLSLQGRSEKGTLDANLYQNNYENFLSEBQKLTTSQDVCTQMNYYGWCSPYSEK 598  
 Db 539 THTLSLQGRSEKGTLDANLYQNNYENFLSEBQKLTTSQDVCTQMNYYGWCSPYSEK 598  
 QY 541 THTLSLQGRSEKGTLDANLYQNNYENFLSEBQKLTTSQDVCTQMNYYGWCSPYSEK 600  
 Db 541 THTLSLQGRSEKGTLDANLYQNNYENFLSEBQKLTTSQDVCTQMNYYGWCSPYSEK 600  
 QY 599 PEWQONIDKARIRGIELTGRNLVTVKASFPVEGKLFGLSLGYAKSKLSDGNSLLSTQPP 658  
 Db 599 PEWQONIDKARIRGIELTGRNLVTVKASFPVEGKLFGLSLGYAKSKLSDGNSLLSTQPP 658  
 QY 601 LEWQONIDKARIRGIELTGRNLVTVKASFPVEGKLFGLSLGYAKSKLSDGNSLLSTQPL 660  
 Db 601 LEWQONIDKARIRGIELTGRNLVTVKASFPVEGKLFGLSLGYAKSKLSDGNSLLSTQPL 660  
 QY 659 KVIAGVDYSPSEKMGVFSRLTYLGAKKAKDAQYTYVYENKGRGTPLQKKVDPYMLNKSAY 718  
 Db 659 KVIAGVDYSPSEKMGVFSRLTYLGAKKAKDAQYTYVYENKGRGTPLQKKVDPYMLNKSAY 718  
 QY 719 YVPDMYGFYKAKNLTLAGVNVNENKYYTWDLSRLGYSYSTTNAVDRDGLDRYRAG 778  
 Db 719 YVPDMYGFYKAKNLTLAGVNVNENKYYTWDLSRLGYSYSTTNAVDRDGLDRYRAG 778  
 QY 779 GRNYAVSLDWKF 790  
 Db 779 GRNYAVSLEWKF 792

## RESULT 4

AAR95566

ID AAR95566 standard; protein; 791 AA.

```

XX SQ Sequence 790 AA;
Query Match 100.0%; Score 4200; DB 2; Length 790;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 790; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKPLHMLPIAALVGSIFGNPVLAADEAATETTPVKAIEKEVRVKDQLNAPATVERVNLGR 60
DB 1 MKPLHMLPIAALVGSIFGNPVLAADEAATETTPVKAIEKEVRVKDQLNAPATVERVNLGR 60

QY 61 IQQMIRNDKDLVRYSTDVGLSDSGRHOKGFAVRGVEGNRVGVSIDGVSPLDSEENSLYA 120
DB 61 IQQMIRNDKDLVRYSTDVGLSDSGRHOKGFAVRGVEGNRVGVSIDGVSPLDSEENSLYA 120

QY 121 RYGNFNSSRLSIDPELVNRIETAKGDSFNTGSGALGGGVNYQTQGHDLILLDDRRQFGVM 180
DB 121 RYGNFNSSRLSIDPELVNRIETAKGDSFNTGSGALGGGVNYQTQGHDLILLDDRRQFGVM 180

QY 181 MKNGYSSRNREWTNTLGFVSNDRVDAALLYSQRRGHETESAGERGYPVEGAGSGAIIIRG 240
DB 181 MKNGYSSRNREWTNTLGFVSNDRVDAALLYSQRRGHETESAGERGYPVEGAGSGAIIIRG 240

QY 241 SSRGIPDPSKHKYHNFGLKIAQINDKHRIGPSFNGQQGHNTTIBESYNLTASSWREADD 300
DB 241 SSRGIPDPSKHKYHNFGLKIAQINDKHRIGPSFNGQQGHNTTIBESYNLTASSWREADD 300

QY 301 VNRRNANLFYEWTPDPSNWLSSLKADFDYQTTKVAAVNNKGSFPPTDYSTWTNRYNKKOLE 360
DB 301 VNRRNANLFYEWTPDPSNWLSSLKADFDYQTTKVAAVNNKGSFPPTDYSTWTNRYNKKOLE 360

QY 361 NIYNRSMDFRFRFTLRWDSQPLQGGQHRLSLKTFFASRRREFENLRDDYFESVRSRTT 420
DB 361 NIYNRSMDFRFRFTLRWDSQPLQGGQHRLSLKTFFASRRREFENLRDDYFESVRSRTT 420

QY 421 SSIQHPVKTNTYGFSLSDQIQWNVFSSRADIRYDHTKMTPOELNAEACHADCKTTPPAANT 480
DB 421 SSIQHPVKTNTYGFSLSDQIQWNVFSSRADIRYDHTKMTPOELNAEACHADCKTTPPAANT 480

QY 481 YKQWGFVGLAQLAQNHQVGDITSGYRVPNASEVYFTYHNGSGNWLFPNPLKAERSTT 540
DB 481 YKQWGFVGLAQLAQNHQVGDITSGYRVPNASEVYFTYHNGSGNWLFPNPLKAERSTT 540

QY 541 HTLSLQGRSEKGTLDANLYQNNYRNFSLBEOQLITSGDVGCTOMNYYYGMCNPNYSEKPE 600
DB 541 HTLSLQGRSEKGTLDANLYQNNYRNFSLBEOQLITSGDVGCTOMNYYYGMCNPNYSEKPE 600

QY 601 WQWQNDKARIRGLEITGRNLNVTKVASFVPEGWKLFGSLGYAKSKLSGDSNLSLSTQPPKV 660
DB 601 WQWQNDKARIRGLEITGRNLNVTKVASFVPEGWKLFGSLGYAKSKLSGDSNLSLSTQPPKV 660

QY 661 IAGVDYESPEKGFVSRITLTYLGAKKAKADQATVYENKRGRTFLOKKVDYFWLNKSAIV 720
DB 661 IAGVDYESPEKGFVSRITLTYLGAKKAKADQATVYENKRGRTFLOKKVDYFWLNKSAIV 720

QY 721 FDMYGFYKLAKNLTLAGVYNNYFNRYTTWDSLRGLYSYSTTNVADRDGKGLDRYRASGR 780
DB 721 FDMYGFYKLAKNLTLAGVYNNYFNRYTTWDSLRGLYSYSTTNVADRDGKGLDRYRASGR 780

QY 781 NYAVSLDWKF 790
DB 781 NYAVSLDWKF 790

```

## RESULT 2

AAR95565  
ID AAR95565 standard; protein; 790 AA.

XX  
AC AAR95565;

XX 27-AUG-2003 (revised)  
DT 25-MAR-2003 (revised)

DT 28-AUG-1996 (first entry)

```

XX N. meningitidis serotype C haemoglobin receptor.
XX Bacterial haemoglobin receptor; hmbR gene; antibody; vaccine; meningitis.
XX Neisseria meningitidis serogroup C.
XX WO9612020-A2.
XX 25-APR-1996.
XX 17-OCT-1995; 95WO-US013623.
XX 18-OCT-1994; 94US-00326670.
XX 02-OCT-1995; 95US-00537361.
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX Stojiljkovic I, So M, Hwa V, Heffron F, Naessif X;
XX WPI: 1996-222006/22.
XX N-PSDB; AAT26997.
XX DNA encoding Neisseria haemoglobin receptor proteins - for use in
XX preparing polypeptide(s) and antibodies for vaccines for e.g. meningitis.
XX Claim 2; Page 50-52; 104pp; English.
XX A bacterial haemoglobin receptor (AAR95565) was identified as the product
XX of the hmbR gene (AAR26997) of N. meningitidis serotype C isolate 8013
XX clone 6. The haemoglobin receptor is thought to be involved in hemin
XX utilisation by the bacterium. This dependence on host iron stores is a
XX potential route for therapeutic intervention strategies. The receptor can
XX be expressed in transformed hosts, e.g. attenuated Salmonella cells. It
XX is useful as a vaccine for meningitis, or can be used to raise antibodies
XX of diagnostic appln. Similar receptor proteins were obtd. from N.
XX meningitidis serotypes A and B and from N. gonorrhoeae (see also AAR95566
XX -68). (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-
XX 2003 to correct PI field.) (Updated on 27-AUG-2003 to correct OS field.)
XX SQ Sequence 790 AA;
Query Match 90.8%; Score 3814; DB 2; Length 790;
Best Local Similarity 91.0%; Pred. No. 5.4e-310;
Matches 720; Conservative 25; Mismatches 44; Indels 2; Gaps 2;

QY 1 MKPLHMLPIAALVGSIFGNPVLAADEAATETTPVKAIEKEVRVKDQLNAPATVERVNLGR 60
DB 1 MKPLHMLPIAALVGSIFGNPVLAADEAATETTPVKAIEKEVRVKDQLNAPATVERVNLGR 60

QY 61 IQQMIRNDKDLVRYSTDVGLSDSGRHOKGFAVRGVEGNRVGVSIDGVSPLDSEENSLYA 120
DB 61 IQQMIRNDKDLVRYSTDVGLSDSGRHOKGFAVRGVEGNRVGVSIDGVSPLDSEENSLYA 120

QY 121 RYGNFNSSRLSIDPELVNRIETAKGDSFNTGSGALGGGVNYQTQGHDLILLDDRRQFGVM 180
DB 121 RYGNFNSSRLSIDPELVNRIETAKGDSFNTGSGALGGGVNYQTQGHDLILLDDRRQFGVM 180

QY 181 MKNGYSSRNREWTNTLGFVSNDRVDAALLYSQRRGHETESAGERGYPVEGAGSGAIIIRG 240
DB 181 MKNGYSSRNREWTNTLGFVSNDRVDAALLYSQRRGHETESAGERGYPVEGAGSGAIIIRG 240

QY 241 SSRGIPDPSKHKYHNFGLKIAQINDKHRIGPSFNGQQGHNTTIBESYNLTASSWREADD 300
DB 241 SSRGIPDPSKHKYHNFGLKIAQINDKHRIGPSFNGQQGHNTTIBESYNLTASSWREADD 300

QY 301 VNRRNANLFYEWTPDPSNWLSSLKADFDYQTTKVAAVNNKGSFPPTDYSTWTNRYNKKOLE 360
DB 301 VNRRNANLFYEWTPDPSNWLSSLKADFDYQTTKVAAI-NKGSFPPTNYTTWETXHKKEVG 359

QY 361 NIYNRSMDFRFRFTLRWDSQPLQ-L-GGOHRLSLKTFFASRRREFENLRDDYFESVRSRT 419
DB 360 EIYNRSMDFRFRFTLRDLSHPLQLGGGRHRLSFXTFASRRDFENLRDDYFESVRSVRT 419

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2005, 14:49:54 ; Search time 168 Seconds  
(without alignments)  
1818.695 Million cell updates/sec

Title: US-09-665-358-8

Perfect score: 4200

Sequence: 1 MKPLHMLPTAALVGSIFGNP.....GLDRYASGRNVAVSLDWKF 790

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score  | Query Match % | Length | DB ID | Description        |
|------------|--------|---------------|--------|-------|--------------------|
| 1          | 4200   | 100.0         | 790    | 2     | AAR95568 N. gonorr |
| 2          | 3814   | 90.8          | 790    | 2     | AAR95565 N. mening |
| 3          | 3812   | 90.8          | 792    | 2     | AAR95567 N. mening |
| 4          | 3655.5 | 87.3          | 791    | 2     | AAR95566 N. mening |
| 5          | 1462.5 | 34.8          | 288    | 6     | ABP76983 N. gonorr |
| 6          | 1445   | 34.4          | 284    | 6     | ABP77003 N. gonorr |
| 7          | 760    | 18.1          | 1046   | 8     | ADT05707 Haemophil |
| 8          | 755    | 18.0          | 809    | 6     | ABP77288 N. gonorr |
| 9          | 753.5  | 17.9          | 967    | 3     | AB44536 Virulence  |
| 10         | 753.5  | 17.9          | 967    | 5     | ABP54488 Pasteurel |
| 11         | 737.5  | 17.6          | 669    | 3     | AB44565 Virulence  |
| 12         | 737.5  | 17.6          | 669    | 5     | ABP54517 Pasteurel |
| 13         | 702.5  | 16.7          | 789    | 7     | ABO78265 Pseudomon |
| 14         | 669.5  | 15.9          | 943    | 5     | ABG91056 Neisseria |
| 15         | 663.5  | 15.8          | 1046   | 8     | ADT05701 Haemophil |
| 16         | 661.5  | 15.8          | 943    | 8     | ADP08381 Neisseria |
| 17         | 649    | 15.5          | 698    | 6     | ABU49701 Protein e |
| 18         | 642.5  | 15.3          | 971    | 3     | AB21223 Haemophil  |
| 19         | 640    | 15.2          | 931    | 2     | AAW04867 Transferr |
| 20         | 631    | 15.0          | 1070   | 2     | AAW35314 M. catarr |
| 21         | 629.5  | 15.0          | 1074   | 2     | AAW35312 M. catarr |
| 22         | 625.5  | 14.9          | 1074   | 2     | AAW35381 M. catarr |
| 23         | 611    | 14.5          | 116    | 6     | ABP76975 N. gonorr |
| 24         | 606    | 14.4          | 930    | 2     | AAW18061 Pasteurel |
| 25         | 597.5  | 14.2          | 911    | 2     | AAR77892 Bacterial |

#### ALIGNMENTS

##### RESULT 1

AAR95568  
ID AAR95568 standard; protein; 790 AA.

XX AC AAR95568;

XX XX 27-AUG-2003 (revised)

DT DT 25-MAR-2003 (revised)

DT DT 28-AUG-1996 (first entry)

XX XX

DE DE N. gonorrhoeae B haemoglobin receptor.

XX XX

KW KW Bacterial haemoglobin receptor; hmbR gene; antibody; vaccine; gonorrhoea.

XX XX

OS Neisseria gonorrhoeae.

XX XX

PN PN WO9612020-A2.

XX XX

PD PD 25-APR-1996.

XX XX

PF PF 17-OCT-1995; 95WO-US013623.

XX XX

PR PR 18-OCT-1994; 94US-00326670.

XX XX

PR PR 02-OCT-1995; 95US-00537361.

XX XX

XX XX (UYOR-) UNIV OREGON HEALTH SCI.

PI PI

XX XX Stojiljkovic I, So M, Hwa V, Heffron F, Nassif X;

XX XX

DR DR WPI; 1996-222006/22.

XX XX

DR DR N-PSDB; AAT27000.

XX XX

PT DNA encoding Neisseria haemoglobin receptor proteins - for use in

XX XX

PT preparing polypeptide(s) and antibodies for vaccines for e.g. meningitis.

XX XX

PS Claim 5; Page 67-69; 104pp; English.

XX XX

CC A bacterial haemoglobin receptor (AAR95568) was identified as the product of the hmbR gene (AAT27000) of N. gonorrhoeae strain MS11A. The haemoglobin receptor is thought to be involved in hemin utilization by the bacterium. This dependence on host iron stores is a potential route for therapeutic intervention strategies. The receptor can be expressed in transformed hosts, e.g. attenuated Salmonella cells. It is useful as a vaccine for gonorrhoea, or can be used to raise antibodies of diagnostic appln. Similar receptor proteins were obtd. from N. meningitidis serotypes A (AAR95566), B (AAR95567) and C (AAR95565). (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to correct OS field.)

CC CC